Monte Carlo Tools for the LHC

LHC@BNL, July 12, 2010

John Campbell, Fermilab

Outline

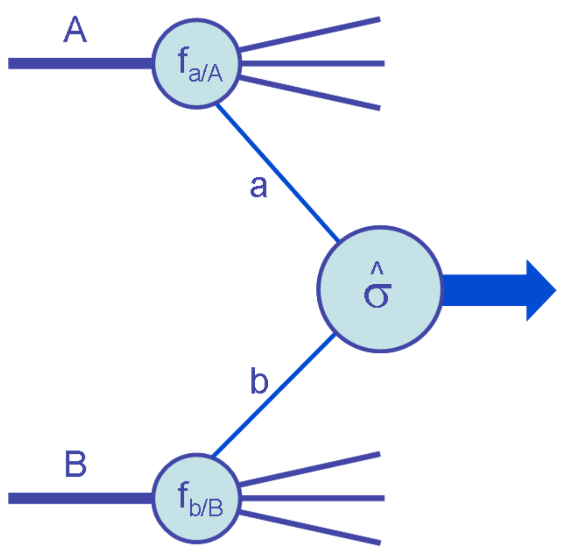
Overview of Monte Carlo tools

MCFM: introduction and practical advice

MCFM: latest updates and physics examples

The keystone: factorization

- Factorization is the foundation of all Monte Carlo descriptions.
- The hadronic cross section factorizes into a part describing the partons inside hadrons (universal) and another part describing the scattering of those partons (calculated case-by-case).



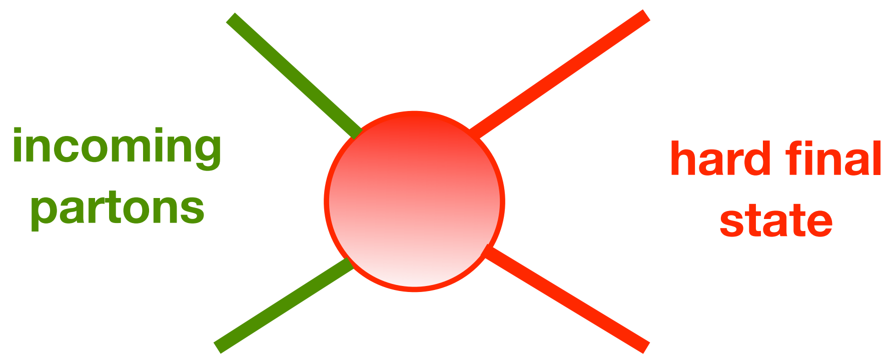
$$\sigma_{AB} = \int dx_a dx_b \; f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \; \hat{\sigma}_{ab \to X}$$

Any hard scale (Q^2) will do, e.g. particle with large enough mass (Drell-Yan) or high E_T object (inclusive jets).

- The presence of this scale means that the hard-scattering cross section may be calculated as a perturbative expansion in the strong coupling, because of asymptotic freedom.
- Typically α_s because plays the dominant role in higher order corrections, although electroweak effects can be important for precision ($\alpha_{ew} \sim \alpha_s^2$).

The hard scattering

- In this talk I will focus on the calculation of the hard scattering.
- Simplest approach: identify the hard final state of interest and compute the cross section from the simplest relevant diagrams.



leading order/tree level
(not always the same thing: loop induced processes, e.g. gg→H)

- A solved problem: cross sections for most processes of interest available from multiple codes.

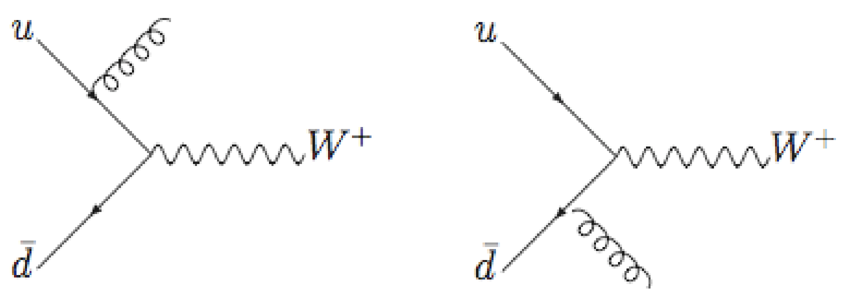
MADGRAPH/MADVENT COMPHEP

WHIZARD HELAC

- The hard is important: without it, the final state is not well defined and the question is meaningless.

Example

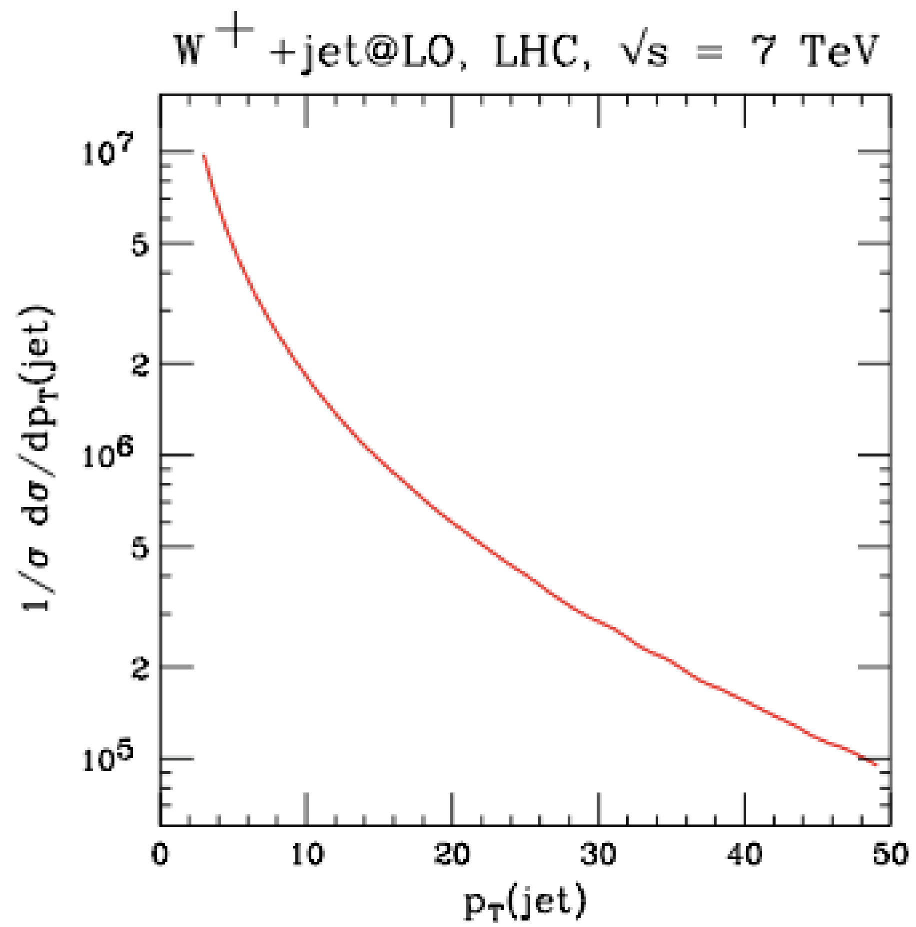
- Final state: W+gluon.



$$|\mathcal{M}^{u\bar{d}\to W+g}|^2 \sim g^2 \left(\frac{\hat{t}^2 + \hat{u}^2 + 2Q^2\,\hat{s}}{\hat{t}\hat{u}} \right),$$

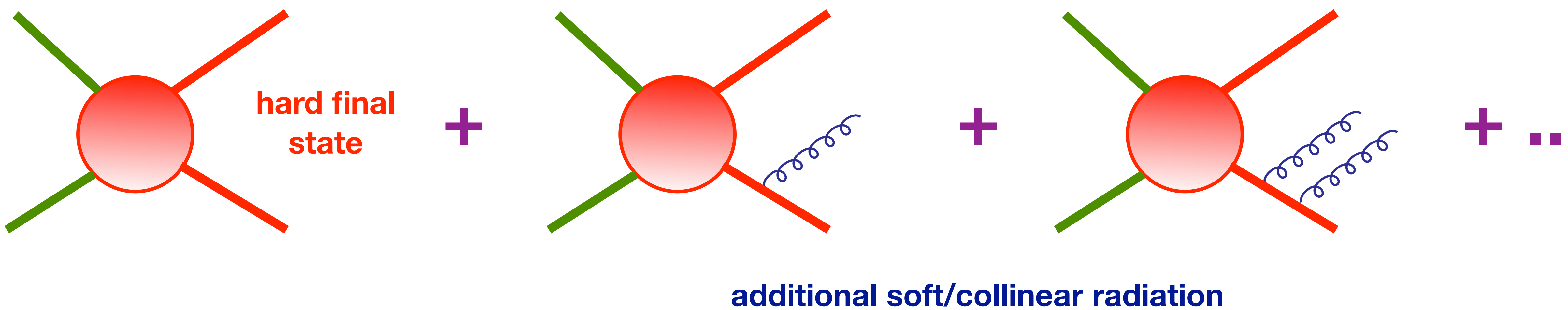
$$\boxed{\hat{s} = s_{u\bar{d}},\ \hat{t} = s_{ug},\ \hat{u} = s_{\bar{d}g}}$$

- There are two types of singularity in the final term:
  - the gluon becomes collinear with the up-quark or anti-d-quark (either $\hat{t}$ or $\hat{u}$ goes to zero on its own)
  - the gluon is soft (both go to zero at the same time).
- By asking for a jet above a given p_T and at finite rapidity we avoid both of these singularities.



Parton showers

- The soft and collinear structure is universal in QCD, and can be exploited to easily generate arbitrary soft and collinear radiation on top of a hard scatter.

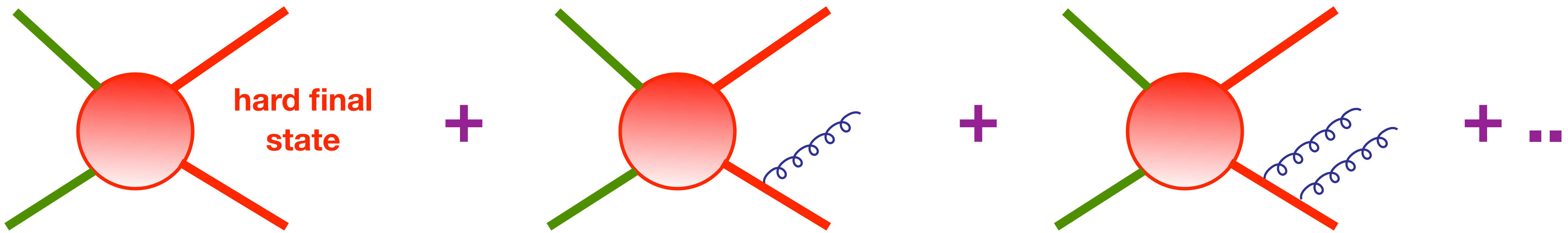


PYTHIA, HERWIG

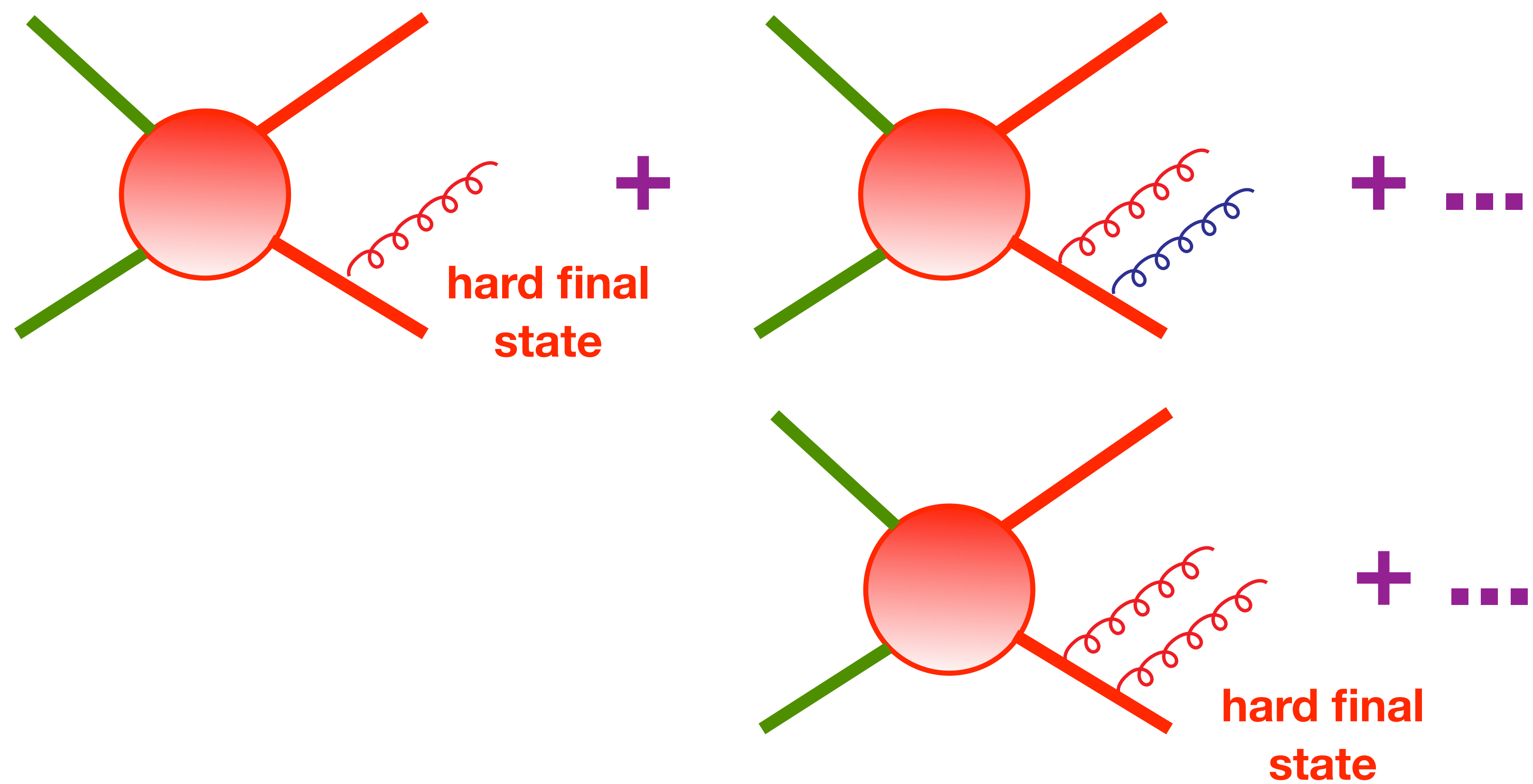
- Starting from a (typically 2→2) hard final state, one can generate large sets of exclusive events similar to those anticipated in the detector.
- The drawback is the approximation of soft and collinear radiation. It is highly favoured, but can miss important kinematic effects.

Improved parton showers

- The solution is to include more hard matrix elements as initial hard scatters, with the trick being to avoid double counting: ME matching/merging.

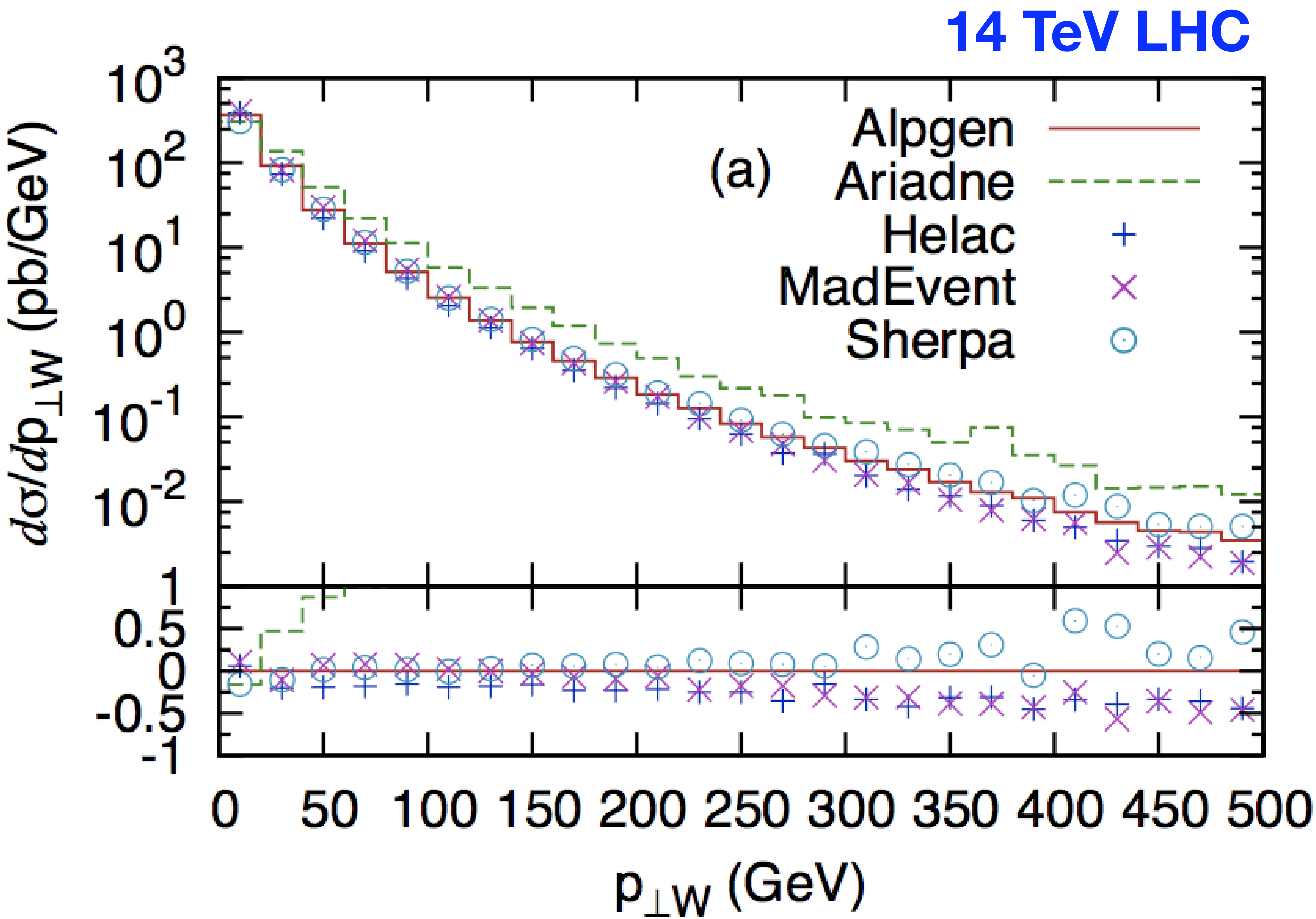


- A number of different schemes for performing the merging exist, e.g. MLM, CKKW, ME&TS (“matrix element +truncated shower”)



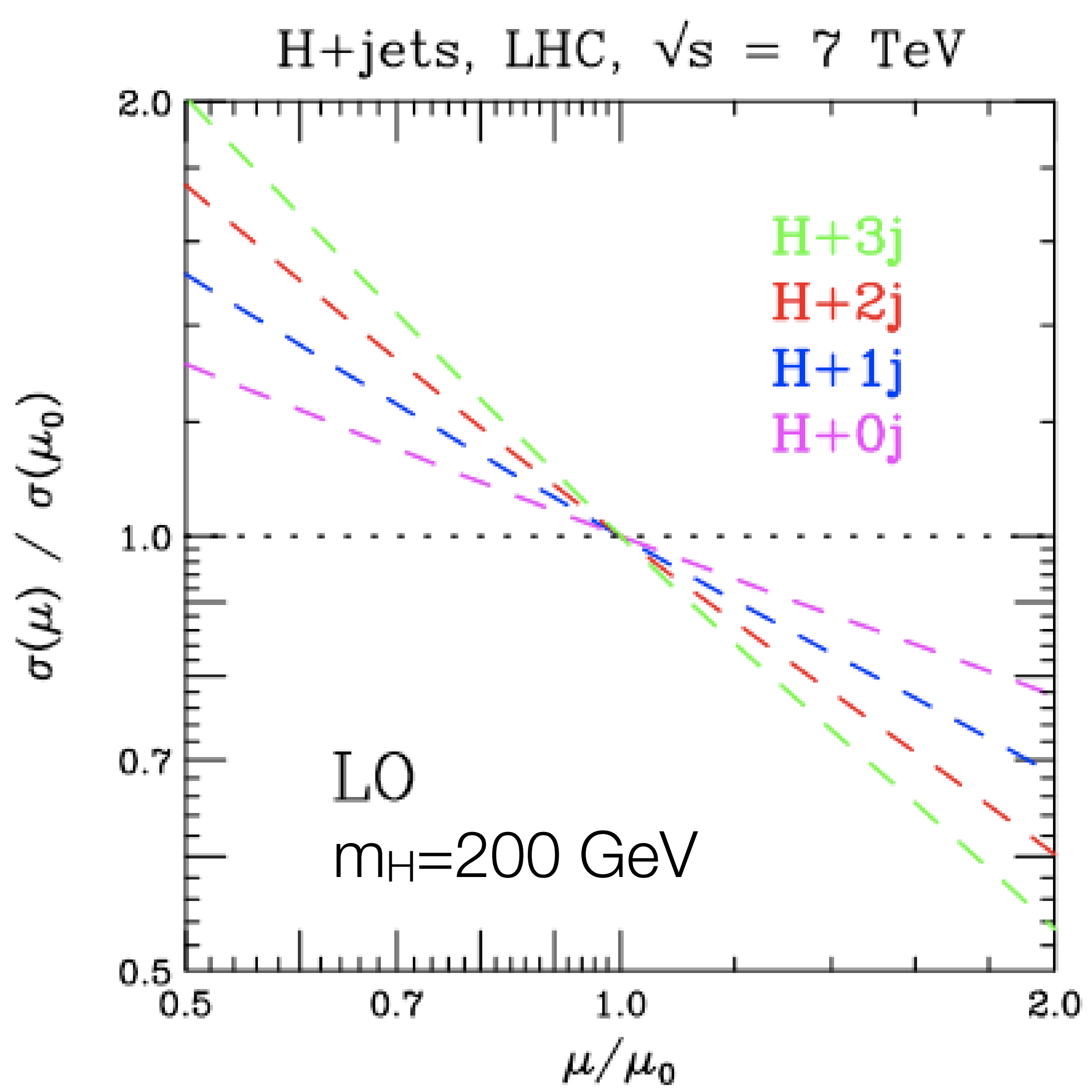
ALPGEN, SHERPA

Comparison of merging techniques

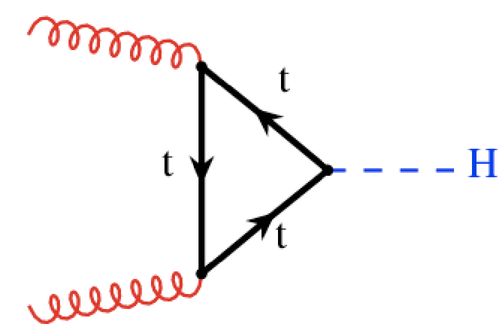


Limits of LO+parton shower

- Despite adding additional radiation with a parton shower, overall normalization of cross section remains a leading order estimate
  - can be very sensitive to renormalization and factorization scales (arguments of strong coupling and PDFs), particularly for final states containing many jets.



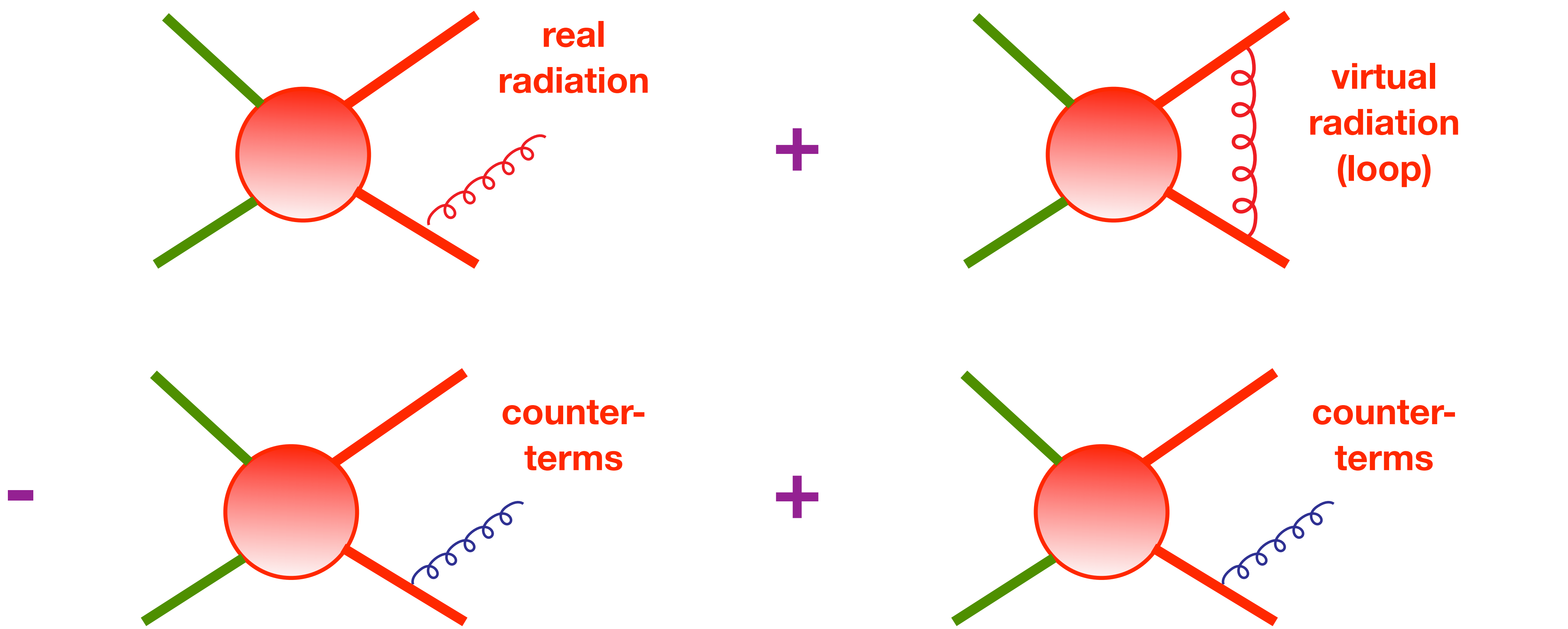
H+jets from gluon fusion



- Prescription (only): theoretical uncertainty ~ scale variation.
- Variation by a factor of two gives δ_{LO}(H+n jets) ≈ 0.25 x (n+1).
- Not good enough for precision studies; improved by going to next-to-leading order (NLO).

Next-to-leading order

- Must include two different contributions and handle soft/coll. singularities.

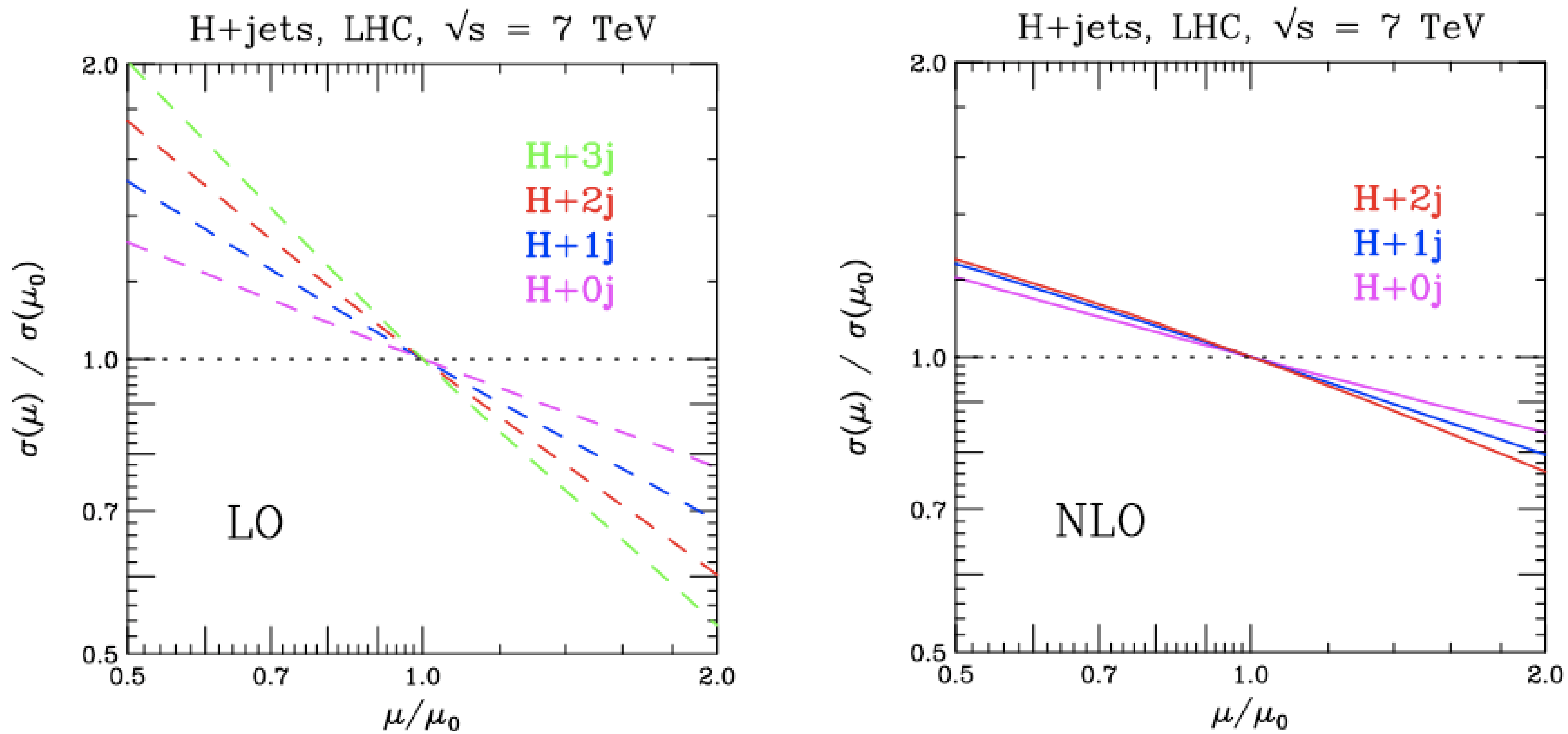


soft/collinear singularities cancelled numerically

singularities cancelled analytically

Mcfm, Nlojet++, Blackhat, Rocket, Helac-1loop

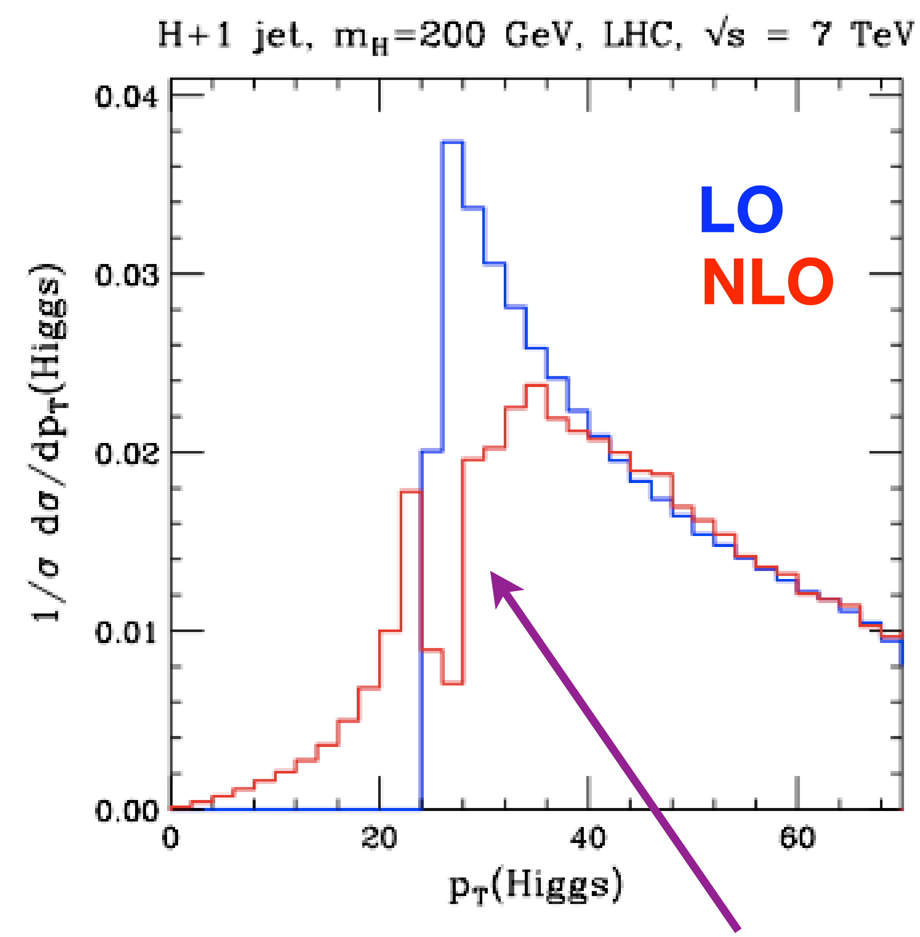
Improved scale dependence



- Reduction in H+0 jet bin relatively small; this results from sensitivity to quark and anti-quark initiated channels that are not present at LO.
- Scale uncertainty reduced to less than 30% in 0-, 1- and 2-jet bins ("lucky" - it did not have to be this good).

Other features

- Compared to LO (without a shower) additional benefits include:
  - exposure to wider range of initial states;
  - sensitivity to final state features such as details of jet algorithm;
  - extended kinematic range.
- Major disadvantages:
  - while calculating LO cross sections is a solved problem, only very recently have we had NLO calculations beyond 2→3 processes.
  - without using a shower, no exclusive hadron-level predictions (just partons).



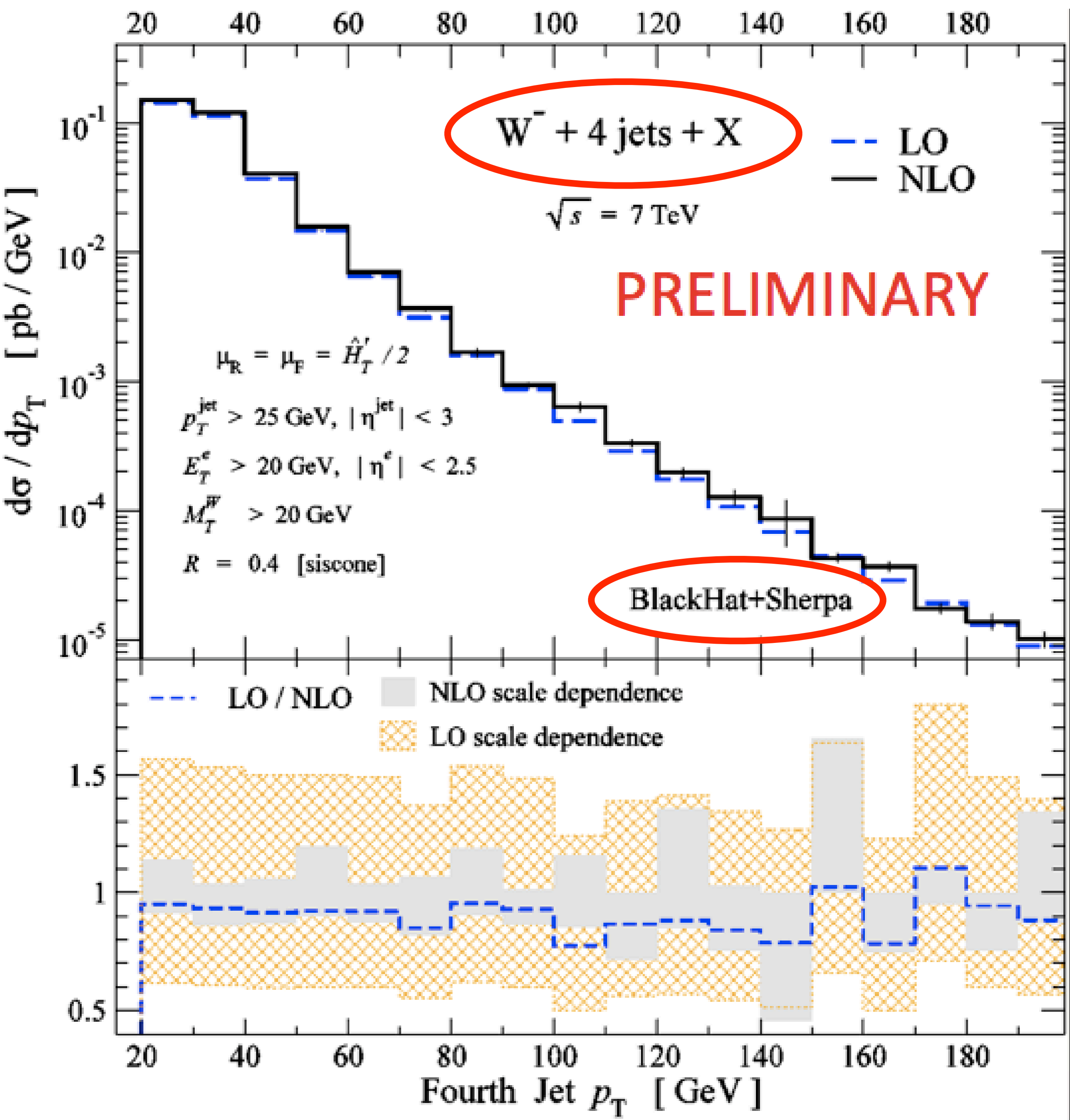
This is an artefact of fixed order
→ prediction unreliable in this region.

At LO the Higgs acquires a p_T by balancing a hard jet (above 25 GeV).

At NLO there are real events where the Higgs balances against two partons with $p_T(j_1)$ and $p_T(j_2) > 25$ GeV but $|p_T(j_1)+p_T(j_2)| < 25$ GeV.

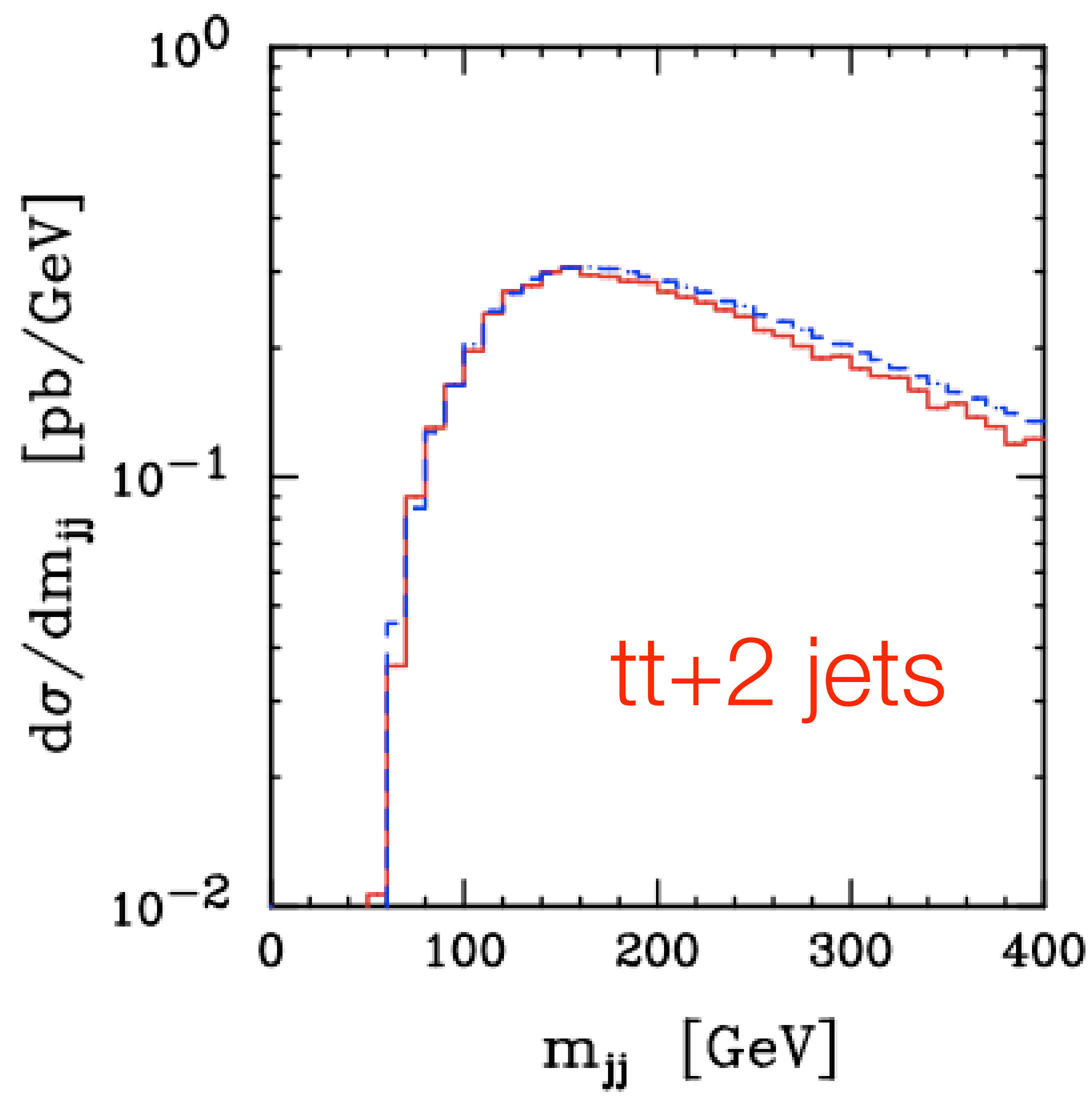
Recent NLO feats

- Algorithmic and numerical computations of NLO corrections are now pushing into regions of jet multiplicity that will be very useful for LHC analyses.



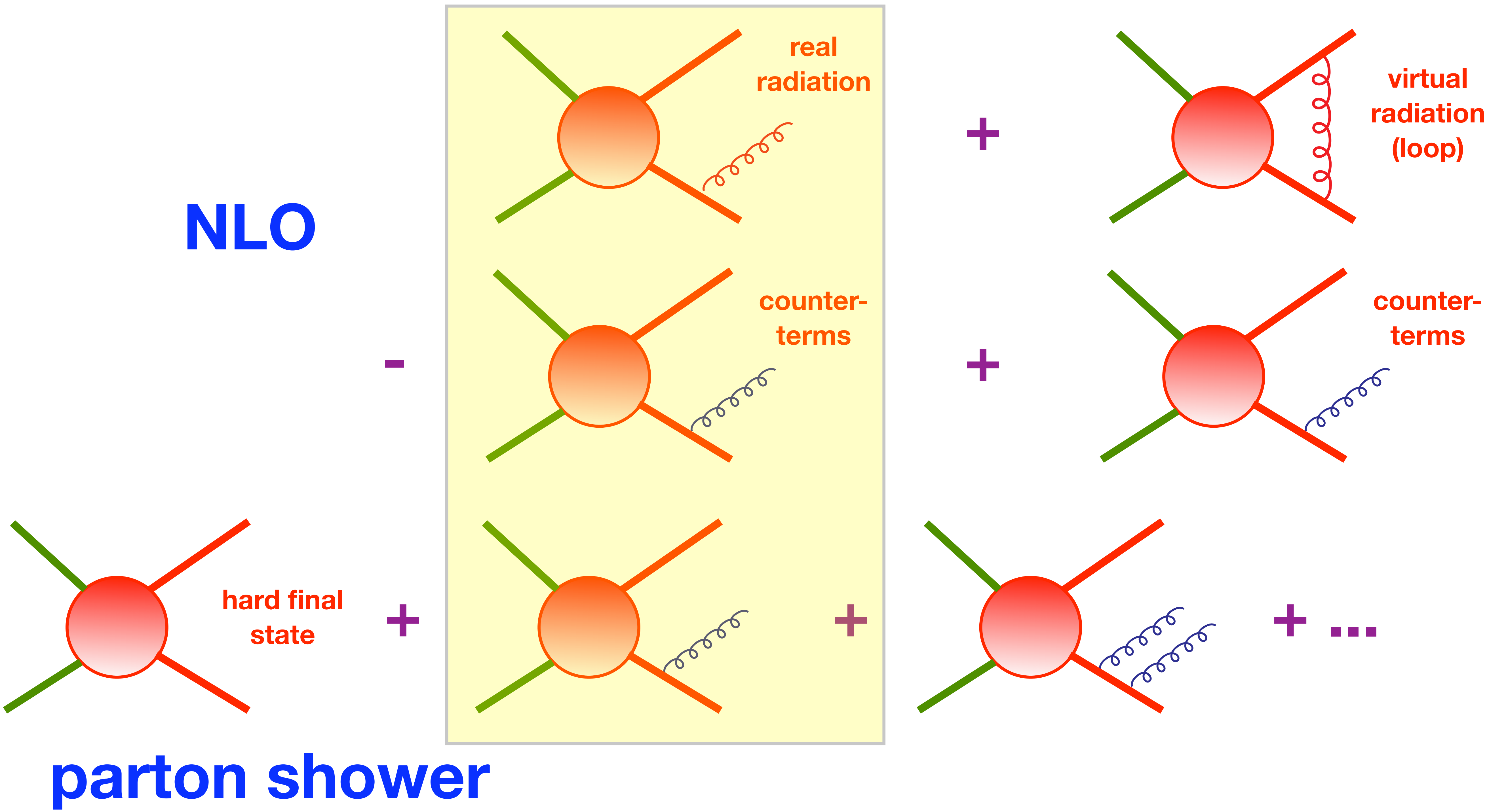
H. Ita (for BlackHat)

Bevilacqua et al. (Helac-1loop)



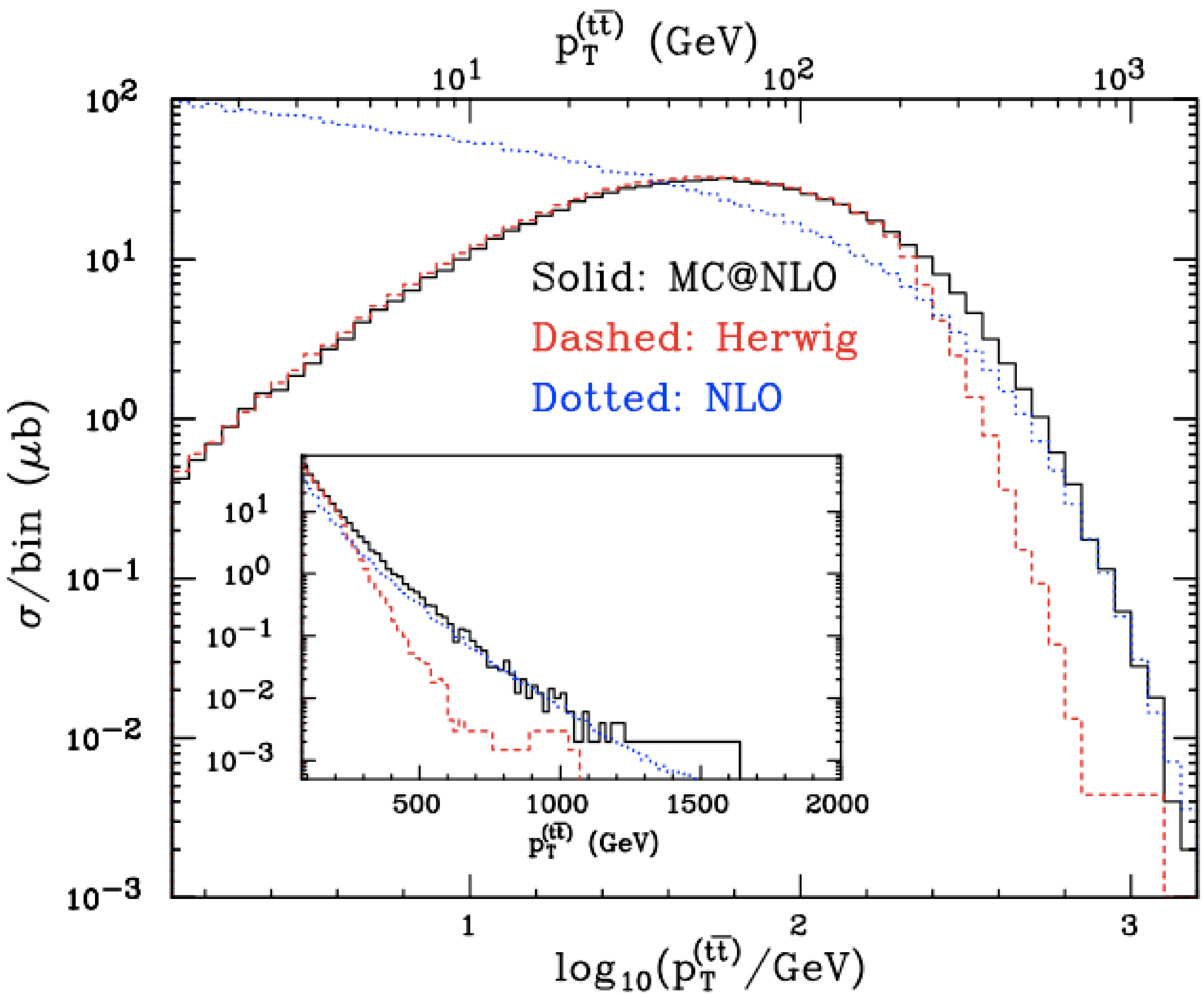
NLO + parton shower

- Major issue is avoiding double-counting in a sensible way.



NLO + PS: MC@NLO

- First real matching of a parton shower (HERWIG) onto a NLO calculation.

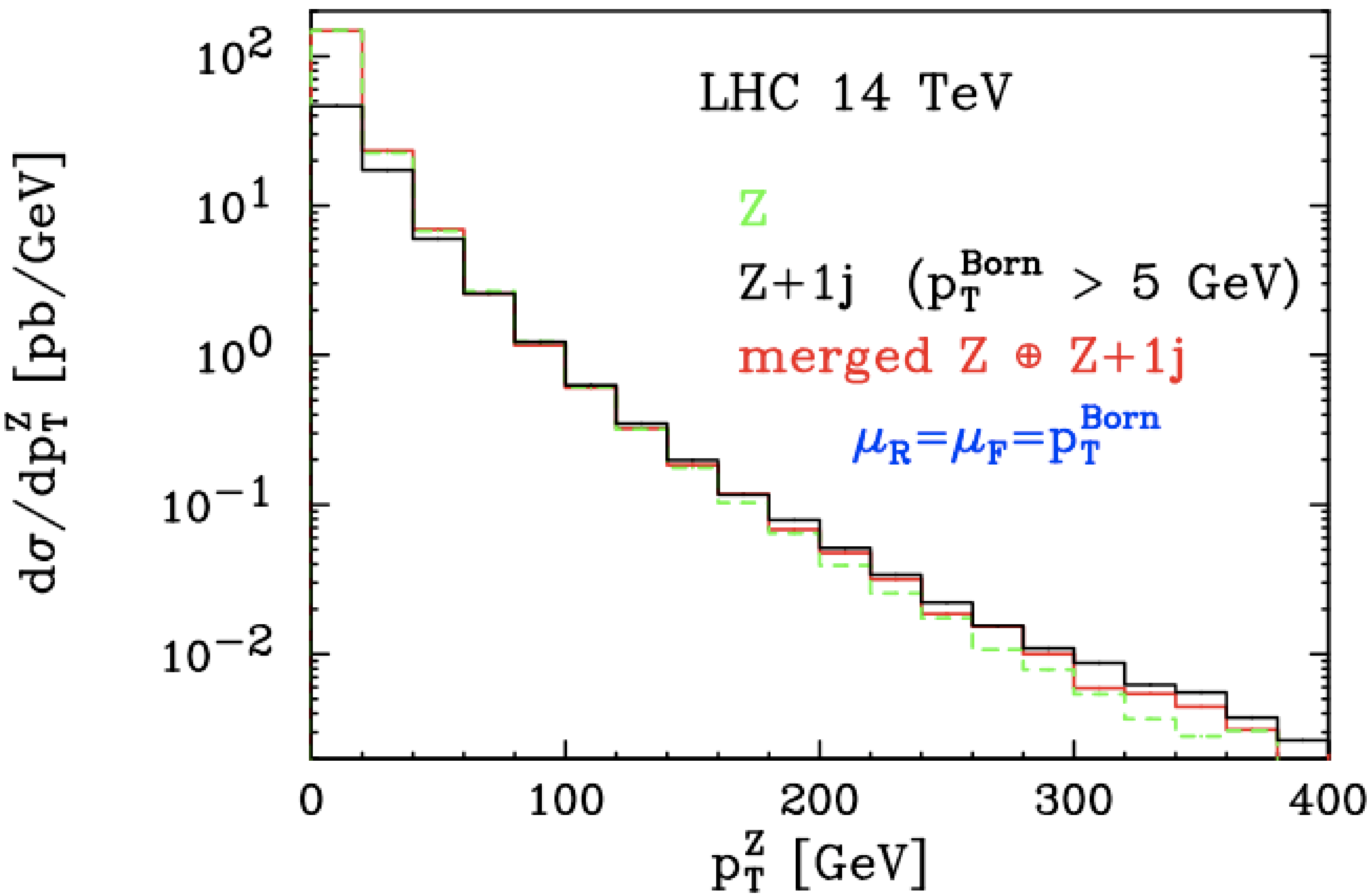


Mc@nlo

Frixione & Webber (2003)

NLO+PS: POWHEG

- More recent implementation, promising simpler procedure through which to "upgrade" parton-level NLO calculations. Your choice of shower.

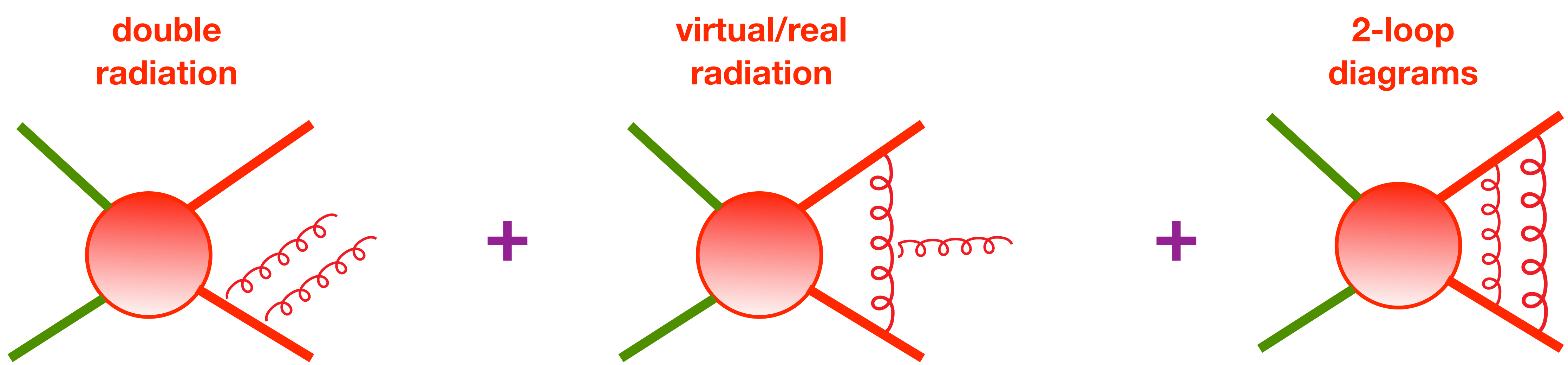


POWHEG: NASON, OLEARI, FRIXIONE ET AL.

FIRST (PRELIMINARY) RESULTS FOR Z+1 JET

NNLO

- For a serious estimate of the error (and precision) need another order.



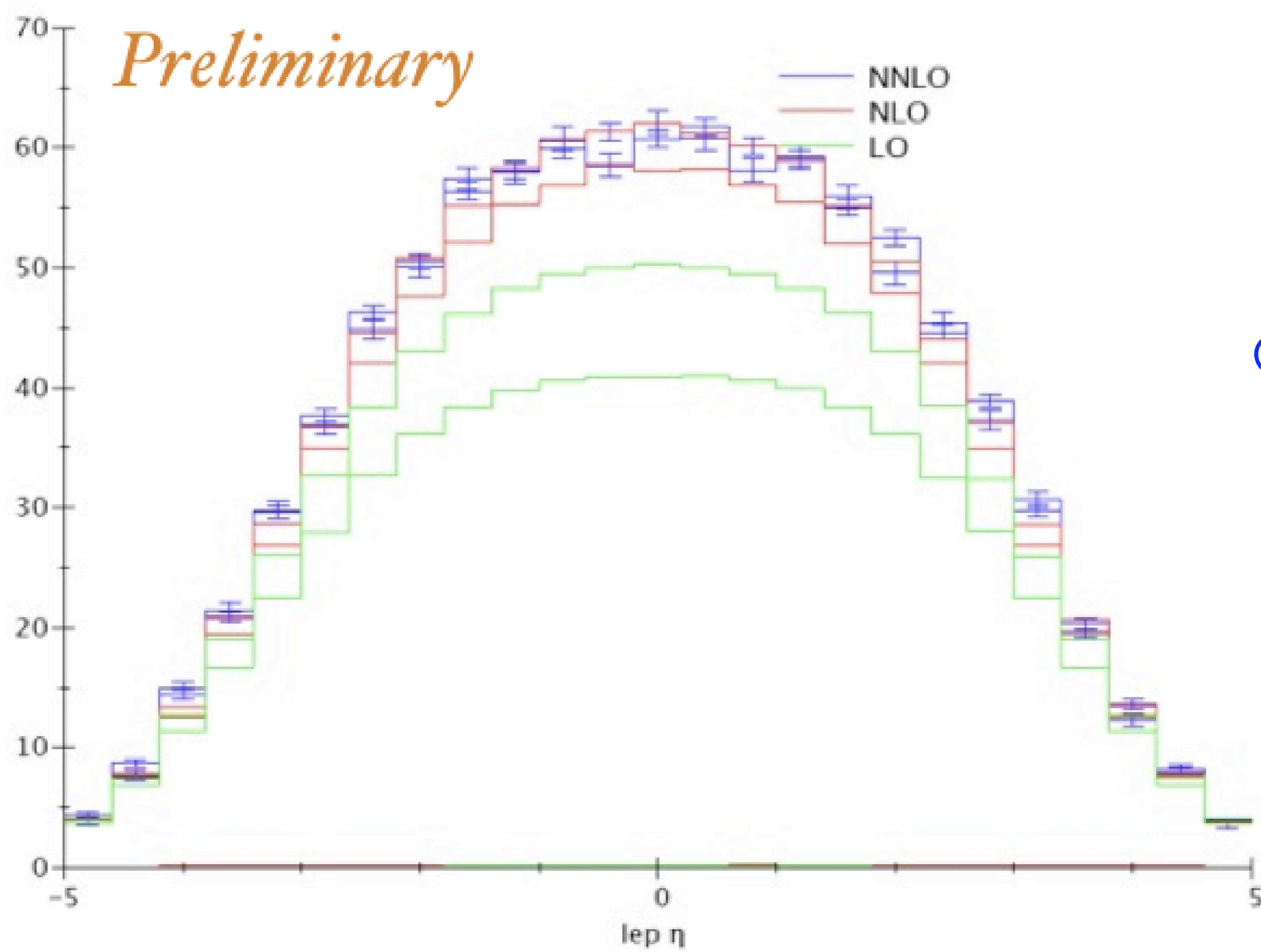
- Two bottlenecks:
  - computing 2-loop integrals for many particles in the final state.
  - organising the cancellation of singularities with appropriate counterterms.
- Very few full calculations exist, e.g. Drell-Yan, Higgs (gluon fusion and WBF).

FEWZ/DYNNLO FEHIP, HNNLO

- No scheme for NNLO + PS at present.

NNLO precision

- Rapidity distibution of leptons in Z production @ 7 TeV.



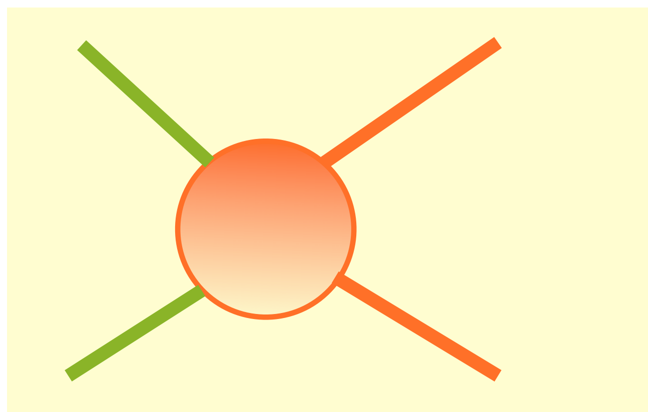
F. Petriello
W/Z workshop
@ BNL, June 2010

- New and improved FEWZ with more efficient phase space integration.

Fixed order recap

- Orders of calculation populate different jet bins at differing orders of accuracy.

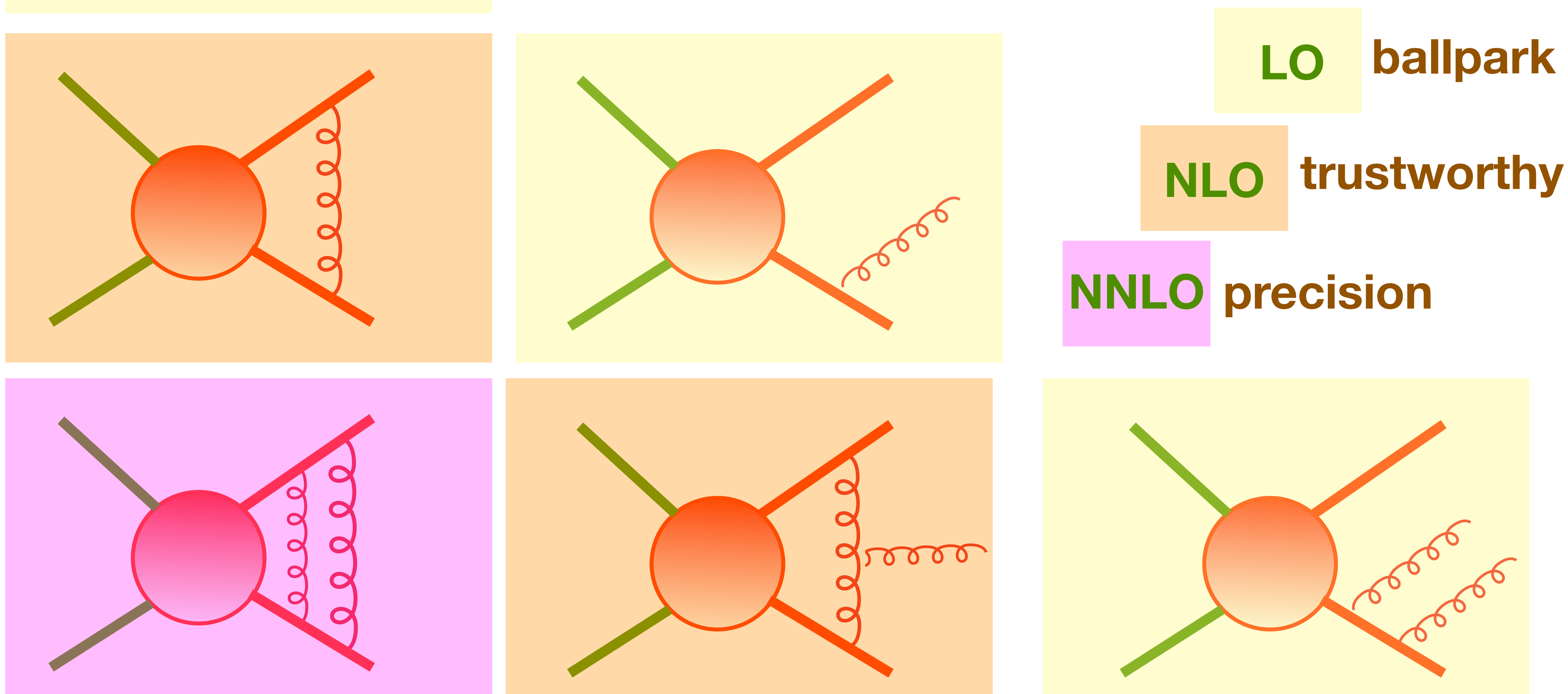
LO N-jet calculation



- When moving beyond normalizing a total cross section, better to think of order of **observable** rather than **calculation**.

NLO N-jet

NNLO N-jet



N-jet kinematics | **(N+1)-jet** | **(N+2)-jet**

MCFM: introduction and practical advice

MCFM summary

- MCFM represents a unified approach to NLO corrections.

  http://mcfm.fnal.gov (v5.8, April 2010)

  J. M. Campbell, R. K. Ellis (main authors)
  R. Frederix, F. Maltoni, F. Tramontano, S. Willenbrock

- Next-to-leading order parton-level predictions.
- Cross sections and differential distributions.
- Standard Model processes involving vector boson+jets, top quarks, Higgs.
- Decays of unstable particles are included, maintaining spin correlations.
- Helicity amplitudes calculated from scratch or taken from the literature.
- Slightly-modified implementation of Catani-Seymour dipole subtraction.

Overview: W/Z+jets

Final state	Notes	Reference
W/Z		
diboson	anomalous couplings	hep-ph/9905386
Wbb	massless b-quark	hep-ph/9810489
Zbb	massless b-quark	hep-ph/0006304
W/Z+1 jet		
W/Z+2 jets		hep-ph/0202176, hep-ph/0308195
Wc	massive c-quark	hep-ph/0506289
Zb	5-flavour scheme	hep-ph/0312024
Zb+jet	5-flavour scheme	hep-ph/0510362

Overview: Higgs and top

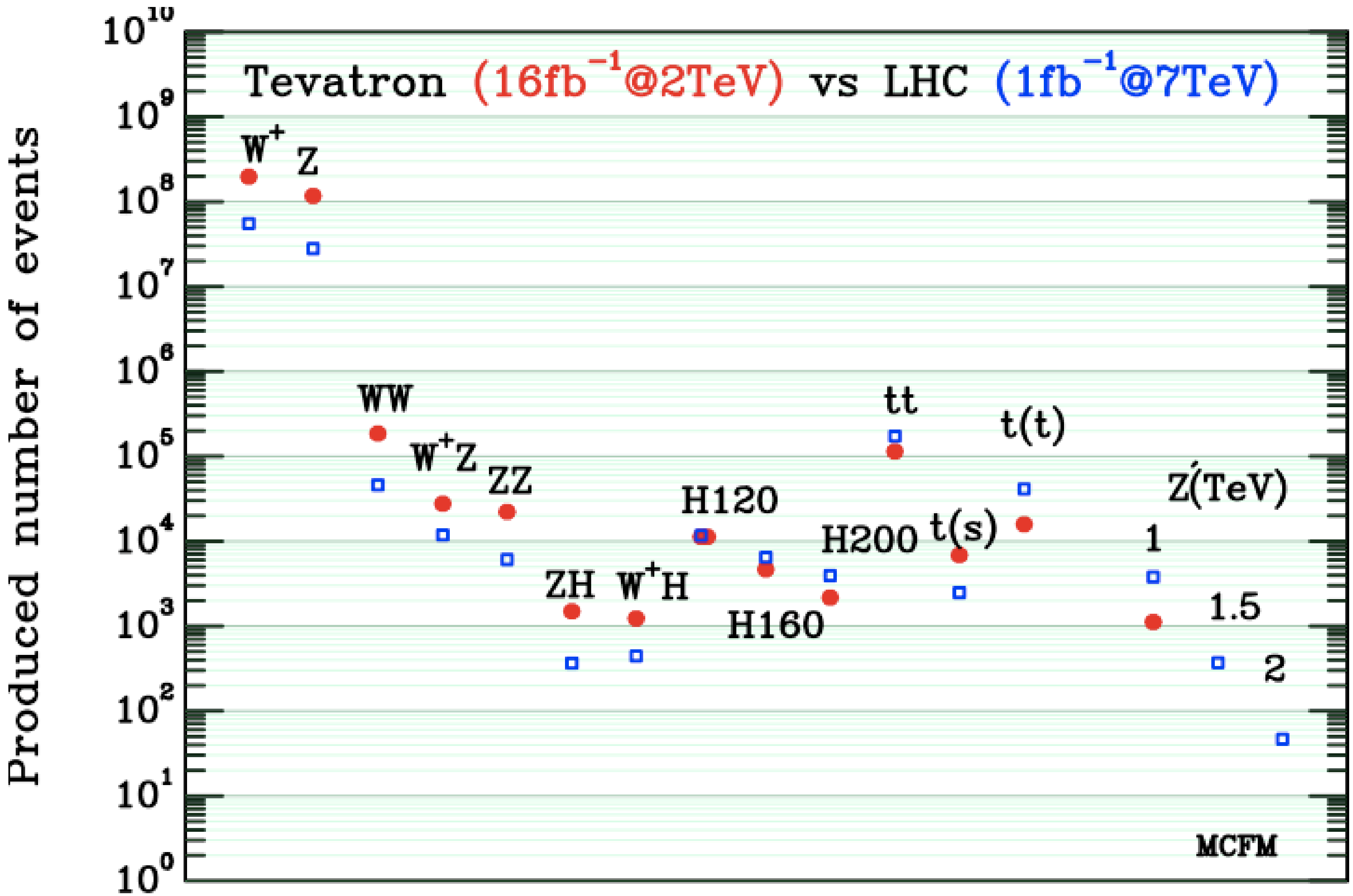
Final state	Notes	Reference
H (g.f.)	effective coupling	
H+1 jet (g.f.)	effective coupling	
H+2 jets (g.f.)	☆ effective coupling	hep-ph/0608194, arXiv:1001.4495
WH/ZH		
H via WBF		hep-ph/0403194
Hb	5-flavour scheme	hep-ph/0204093
t	s- and t-channel (5F), top decay included	hep-ph/0408158
t	t-channel (4F)	arXiv:0903.0005, arXiv:0907.3933
Wt	5-flavour scheme	hep-ph/0506289
top pairs	☆ with top decay	

Overview: other calculations

Final state	Notes	Reference
Wb+jet	complicated procedure, private version only	hep-ph/0611348, arXiv:0809.3003
WW+jet	semi-numerical virtual amplitudes; private	arXiv:0710.1832
J/ψ and Υ(singlet)	private version only, could be made available	hep-ph/0703113, arXiv:0806.3282
J/ψ (photo-prod. in DIS)	private version only, could be made available	arXiv:0901.4352

- Private versions either not sufficiently polished for a general audience, or of only limited interest.
- Leading order calculations of related processes also available.
- Select other processes also at LO (e.g. ttH with decays).

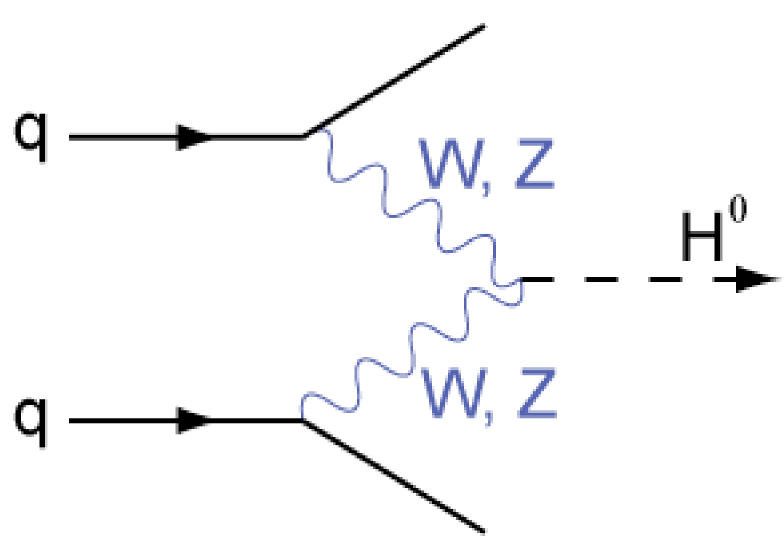
MCFM: expected samples Xmas 2011(... 2014?)



- By the end of 2011, both machines should have similar samples of W, Z, diboson, top and (possibly) low mass Higgs events.
- The LHC wins out for states of higher mass (Higgs, Z', SUSY etc.).

MCFM usage

- The code is controlled by a text file input.DAT
    - choice of basic parameters, masses and simple event cuts;
    - other electroweak inputs specified at compile time.
- Choice of three built-in jet algorithms: cone , k_T and anti-k_T.
- Jet clustering can also be turned off if appropriate (no singularities), e.g. vector boson fusion.



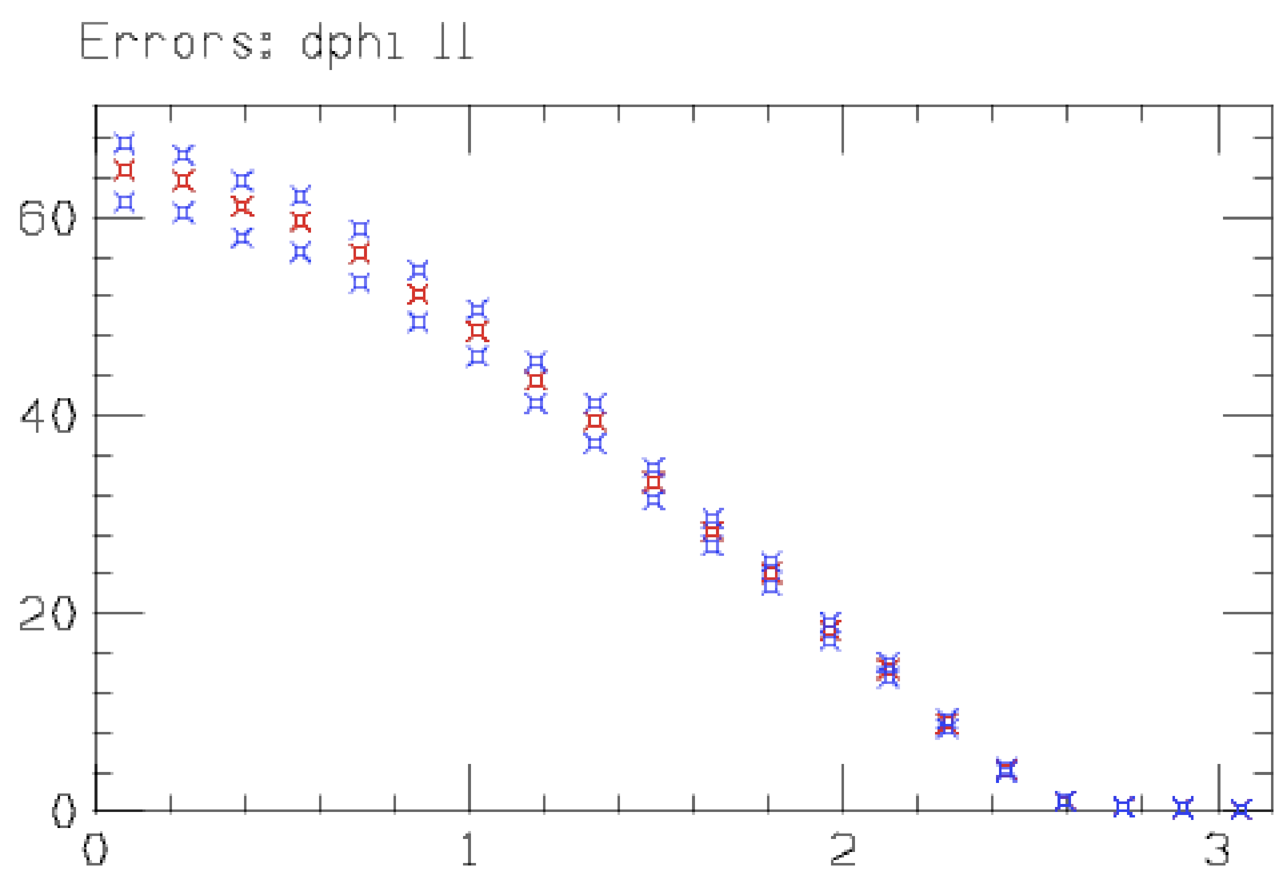

```
'5.7'           [file version number]

[Flags to specify the mode in which MCFM is run]
.false.         [evtgen]
.false.         [creatent]
.false.         [skipnt]
.false.         [dswhisto]

[General options to specify the process and execution]
1               [nproc]
'lord'          [part 'lord','real' or 'virt','tota']
'test'          ['runstring']
14000d0         [sqrts in GeV]
+1              [ih1 =1 for proton and -1 for antiproton]
+1              [ih2 =1 for proton and -1 for antiproton]
120d0           [hmass]
80d0            [scale:QCD scale choice]
80d0            [facscale:QCD fac_scale choice]
.false.         [dynamicscale]
.false.         [zerowidth]
.true.          [removebr]
10              [itmx1, number of iterations for pre-conditioning]
20000           [ncall1]
10              [itmx2, number of iterations for final run]
20000           [ncall2]
1089            [ij]
.false.         [dryrun]
.true.          [Qflag]
.true.          [Gflag]

[Heavy quark masses]
172.5d0         [top mass]
4.75d0          [bottom mass]
1.5d0           [charm mass]

[Pdf selection]
'cteq6l1'       [pdlabel]
4               [NGROUP, see PDFLIB]
46              [NSET - see PDFLIB]
cteq6mE.LHgrid  [LHAPDF group]
-1              [LHAPDF set]

[Jet definition and event cuts]
0d0             [m34min]
14000d0         [m34max]
0d0             [m56min]
14000d0         [m56max]
.true.          [inclusive]
'ktal'          [algorithm]
15d0            [ptjet_min]
0d0             [|etajet|_min]
2d0             [|etajet|_max]
0.7d0           [Rcut_jet]
.false.         [makecuts]
```

PDF features

- Many recent fits implemented natively, but MCFM can also be linked against LHAPDF for the very latest versions and, for example, NNPDF.
- For PDF fits containing uncertainty sets, the PDF uncertainty can be estimated in a single run, e.g. automatically loops over 44 sets in CTEQ6.6.

(asymmetric) uncertainty in the total cross section

```
* --------------- SUMMARY ---------------- *
*                                          *
*           HEPDATA prescription           *
*      (see, for example Eqn. (43) of      *
*       J.Campbell, J.Huston, W.J.Stirling, *
*       Rep. Prog. Phys. 70 (2007) 89)     *
*                                          *
*       Minimum value         94.735 fb    *
*       Central value         97.854 fb    *
*       Maximum value        100.315 fb    *
*                                          *
*    Err estimate +/-          4.714 fb    *
*       +ve direction          4.256 fb    *
*       -ve direction          5.198 fb    *
*    Fractional error          0.048       *
*                                          *
```



PDF uncertainty in a distribution (requires 1-line addition to code)

Output options

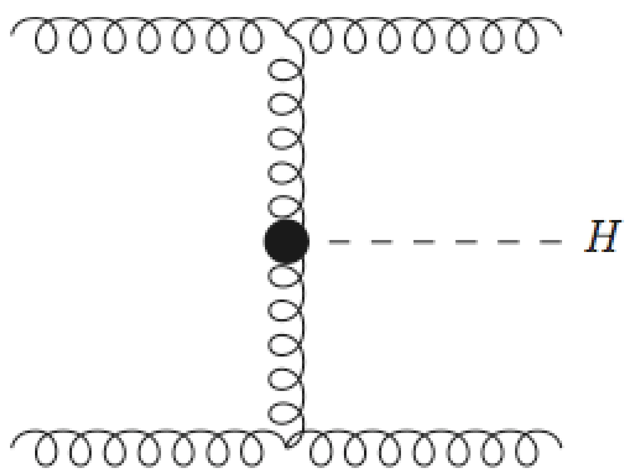
- Default behaviour is to accumulate histograms internally.
  - output to a text data file at the end and (archaic) Topdrawer plots.
- Alternatively, can write event n-tuples directly:
  - ROOT n-tuples are written through the FROOT interface (P. Nadolsky)
  - events are written after jet clustering, so changes to jet definition in processing do not make sense.
  - allows greater flexibility for plotting observables of interest and re-binning.
  - weights for different PDF uncertainty sets are included too.
- In the pipeline ...
  - distribution of a script to convert default Topdrawer histograms into ROOT format (requires ROOT+Python).
  - improved n-tuple support. Pointers linking real and subtracted events, to enable proper statistical analysis.

MCFM: latest updates and physics examples

Higgs + 2 jets

- A new inclusion in the most recent versions of the code has been production of H+2 jets from parton-parton fusion.

not just gluon fusion, also channels with one or two quark pairs



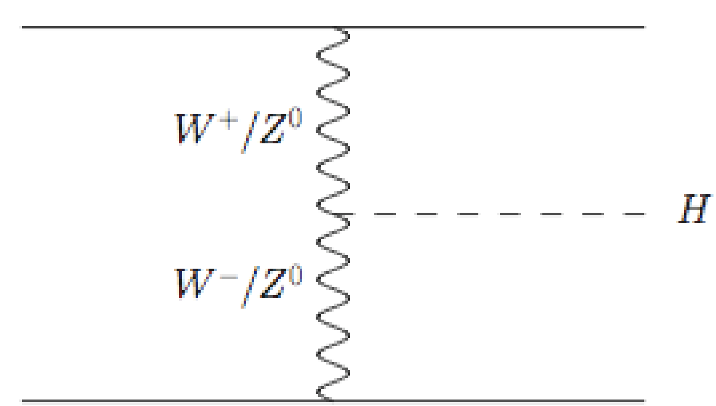
Effective coupling of Higgs to two gluons:

$$\mathcal{L}_{\text{eff}} = \frac{1}{4}\left(\frac{g^2}{12\pi^2 v}\right)(1+\Delta) H G^a_{\mu\nu} G^{a\,\mu\nu}$$

Valid as long as m_H, p_T(jet) $\lesssim m_t$

Del Duca et al. (2001)

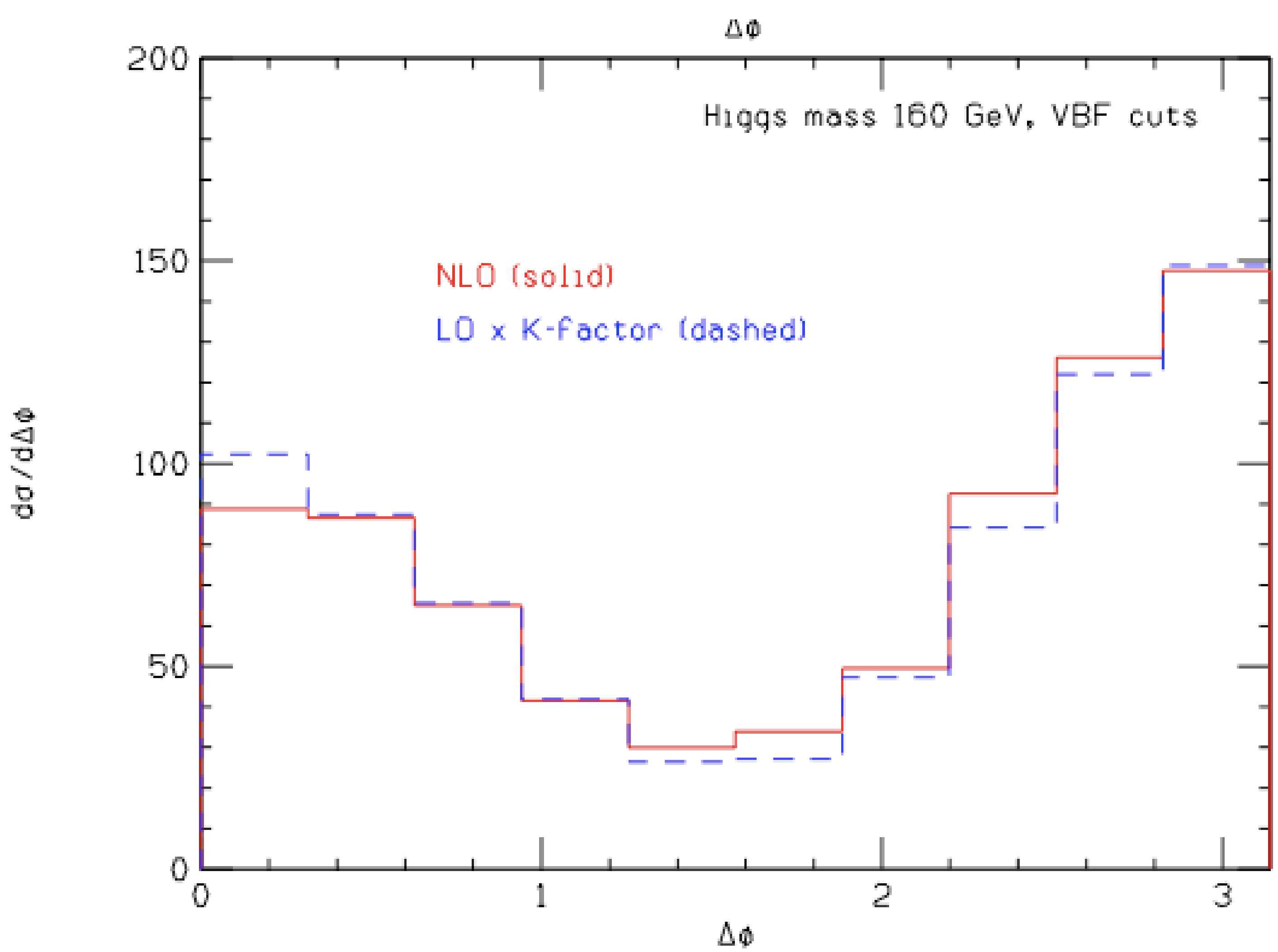
- Interest in this process is high:
  - a 2→3 calculation amenable to a number of calculational techniques;
  - "Every little helps" for the Higgs search at the Tevatron;
  - this is the same final state as for vector boson fusion, relevant for extraction of Higgs couplings @ LHC.



vector boson fusion

First iteration

- First calculation of the relevant matrix elements in 2005.
  Ellis, Giele, Zanderighi
  - 4-quark matrix elements (spin-summed) in analytic form;
  - the remainder (bulk of the calculation) using a semi-numerical method in which tensor integrals are reduced numerically to a known set of scalars.
- A phenomenological study followed shortly after, with limited results due to the intense computational effort required at each phase space point.
  JC, Ellis, Zanderighi (2006)
- As a result, the code was never released.



New implementation

- All helicity amplitudes calculated analytically over last 3-4 years using novel unitarity methods.
    - Implemented in MCFM v.5.7 and onwards.
- Fully checked against the previous semi-numerical implementation.
    - much faster (for comparison: quicker than W/Z+2 jets)
    - evaluation of virtual contribution no longer an issue, so full decays of the Higgs into W pairs can be straightforwardly included.

$$h_1 + h_2 \to H + j_1 + j_2 \to W^- + W^+ + j_1 + j_2$$

$$W^+ \hookrightarrow \nu + e^+$$

$$W^- \hookrightarrow e^- + \bar{\nu}$$

- New phenomenological study: JC, Ellis, Williams, arXiv: 1001.4495.
- Reminder: we're working in the effective coupling limit. At LO, corrections for this approximation can be large for m_H above ~ m_t.

Application: Tevatron Higgs search

- The search strategy at the Tevatron divides events into bins according to the number of jets observed.
  - understanding the expected theoretical cross section and uncertainty in each bin is essential for imposing the best limits.
- A detailed theoretical analysis, using the NNLO calculation of the inclusive Higgs cross section, was given a year ago.

Anastasiou et al. (2009)

$$\frac{\Delta N_{\rm signal}({\rm scale})}{N_{\rm signal}} = \underbrace{60\%}_{\text{0-jet}} \cdot \begin{pmatrix} +5\% \\ -9\% \end{pmatrix} + \underbrace{29\%}_{\text{1-jet}} \cdot \begin{pmatrix} +24\% \\ -23\% \end{pmatrix} + \underbrace{11\%}_{\text{2-jet}} \cdot \begin{pmatrix} +91\% \\ -44\% \end{pmatrix} = \underbrace{\begin{pmatrix} +20.0\% \\ -16.9\% \end{pmatrix}}_{\text{overall}}$$

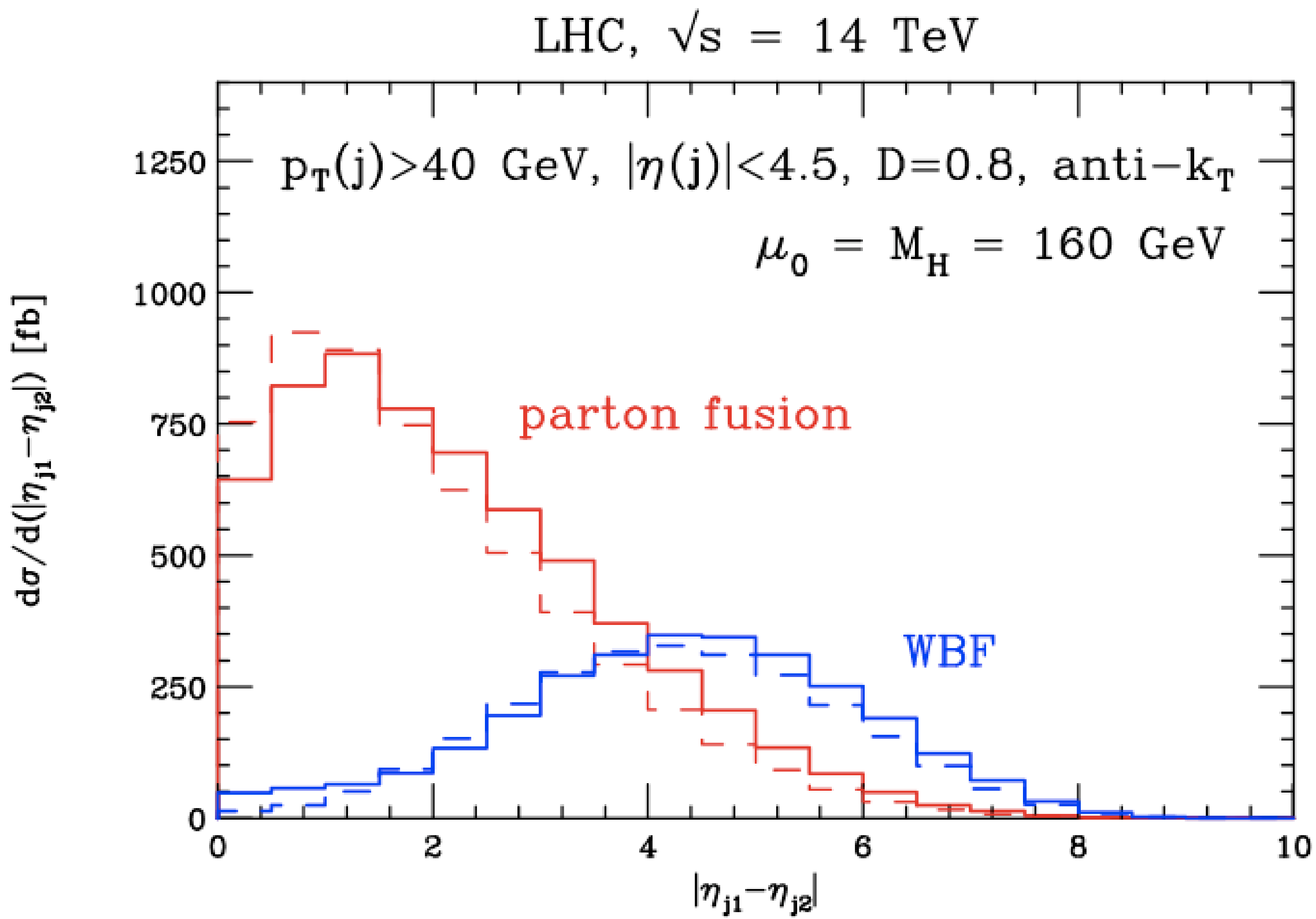
expected jet composition after preselection

- The 2-jet bin is only a small fraction of the total, but the large uncertainty results in a sizeable contribution to the overall uncertainty → update to NLO:

$$\frac{\Delta N_{\rm signal}({\rm scale})}{N_{\rm signal}} = 60\% \cdot \begin{pmatrix} +5\% \\ -9\% \end{pmatrix} + 29\% \cdot \begin{pmatrix} +24\% \\ -23\% \end{pmatrix} + 11\% \cdot \boxed{\begin{pmatrix} +35\% \\ -31\% \end{pmatrix}} = \boxed{\begin{pmatrix} +13.8\% \\ -15.5\% \end{pmatrix}}$$

Comparison with WBF @ 14 TeV LHC

- WBF sample can be isolated by imposing a large rapidity separation between two of the jets, but some “contamination” from parton fusion remains.

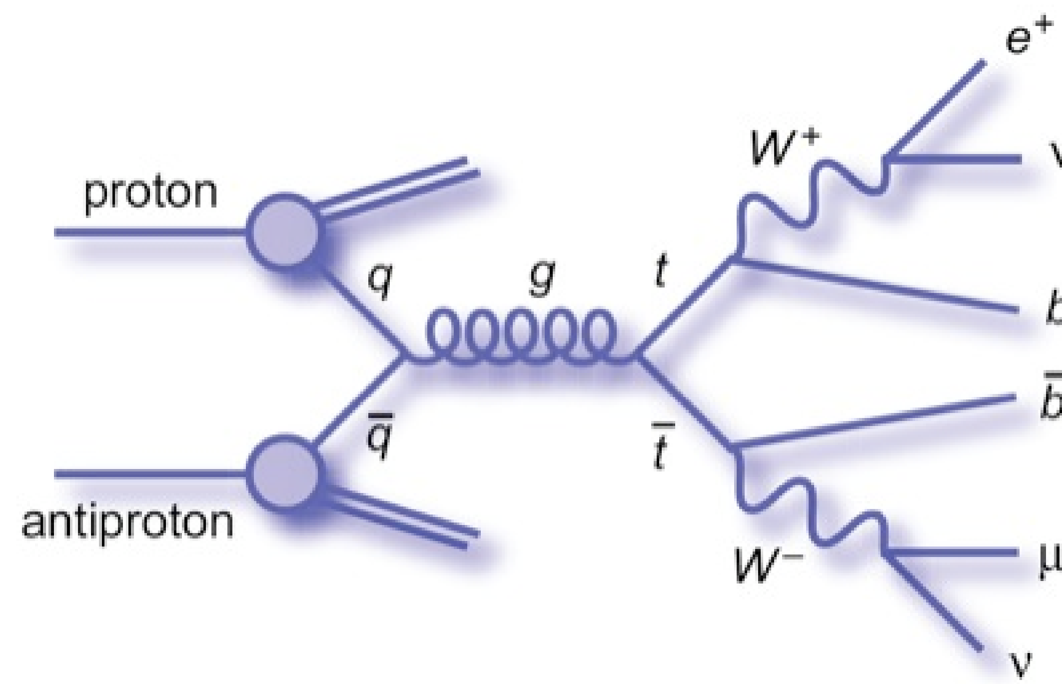


largest absolute rapidity difference between two jets

Top pairs with decay

- Now include full $t \to e^+ \nu b$ decays.
    - spin correlations included; (in next version: can also be turned off, for comparison)
    - top quark must be kept strictly on-shell, no radiation in decay;
    - implementation of a previous matrix element calculation.

  KORNER, MEREBASHVILI (2002)



- This is not a new result - a number of calculations already in the literature.

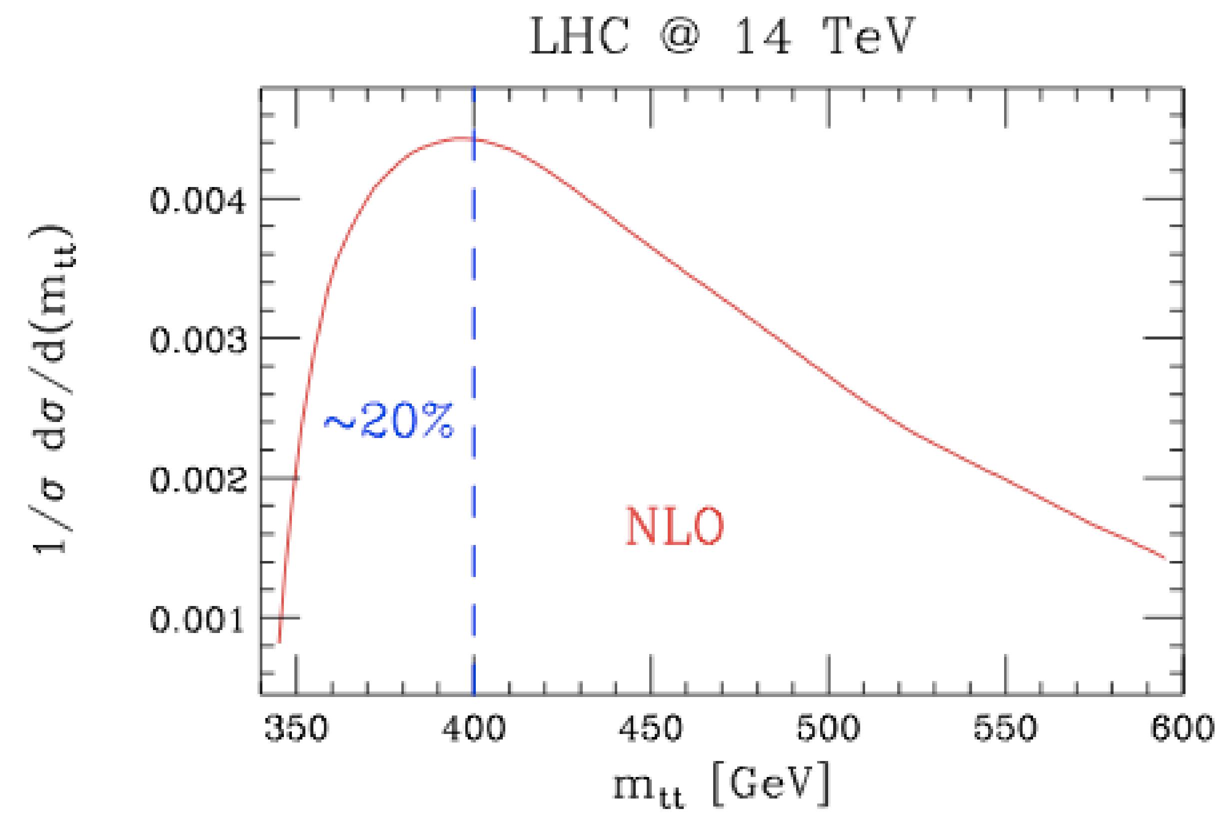
  BERNREUTHER ET AL. (2001)

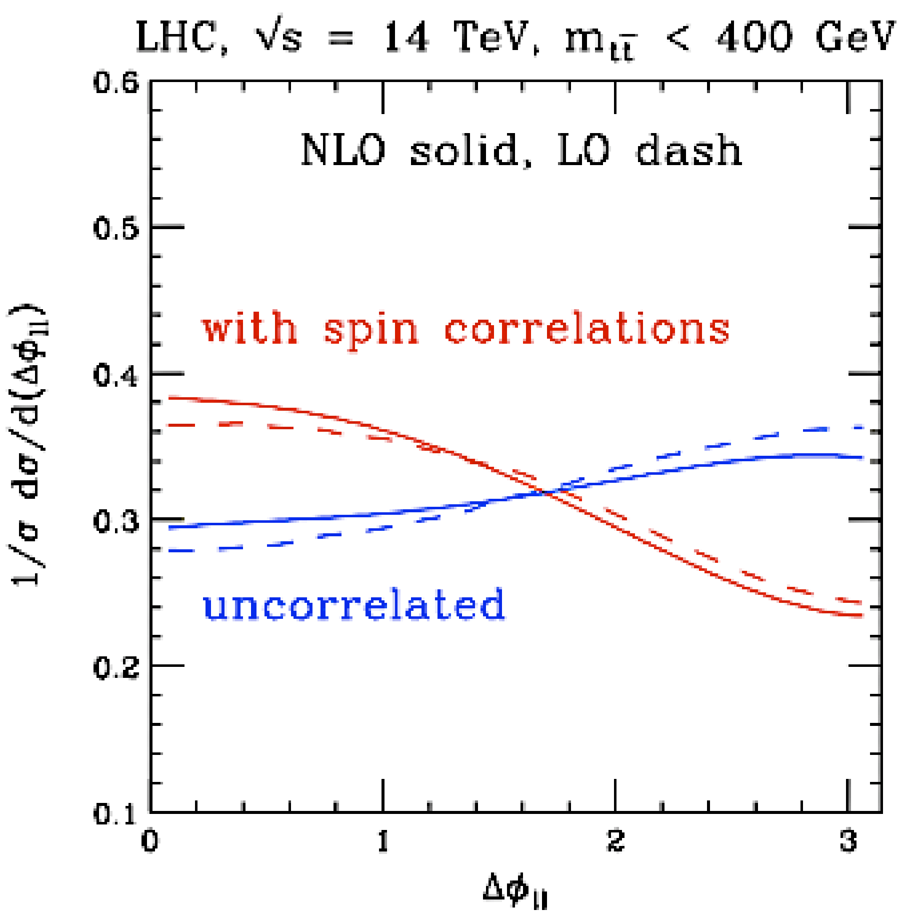
  MELNIKOV, SCHULZE (2009)

- The importance of this process at the LHC - particularly as a background - still merits its inclusion in MCFM (v5.8 onwards).

Measuring spin correlations at the LHC

- Spin correlations in top pair production are more easily observed at the LHC once an invariant mass cut is applied.
  MAHLON, PARKE (2010)
- A cut at 400 GeV still retains about 20% of all top pair events.



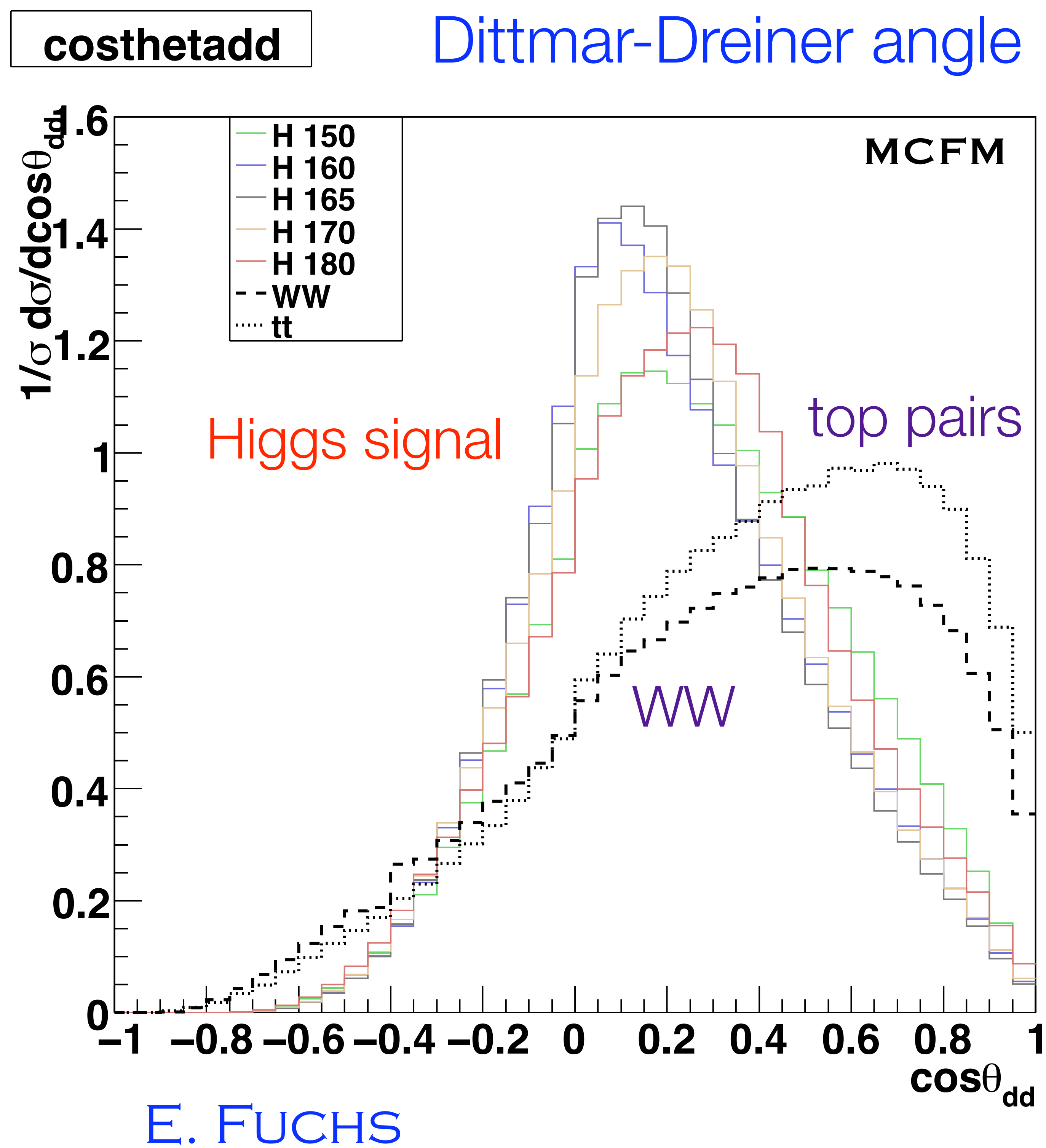


azimuthal angle between leptons

- No. of dilepton events/fb^{-1} ~ 1000 (acceptance+efficiency ~ 10%)
  - plenty for distributions
  - factor of 5 less at 7 TeV, not as convincing (similar to Tevatron)
- Study of the correlation expected at NLO shows slight change from LO.

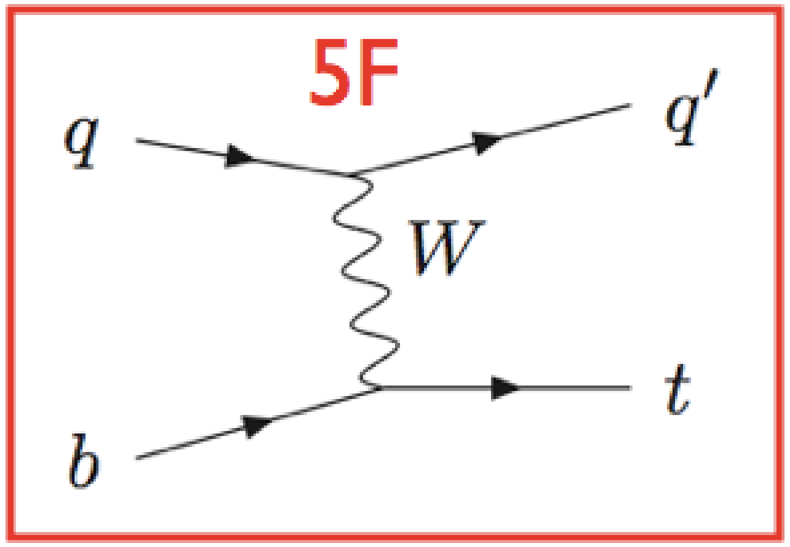
H→WW* search

- Signal and two main backgrounds all included in MCFM, with full spin correlations.
- This plot is part of work performed this summer to try to improve the available documentation for MCFM.
  - survey of physics at 7 TeV.
  - benchmark plots and physics examples.

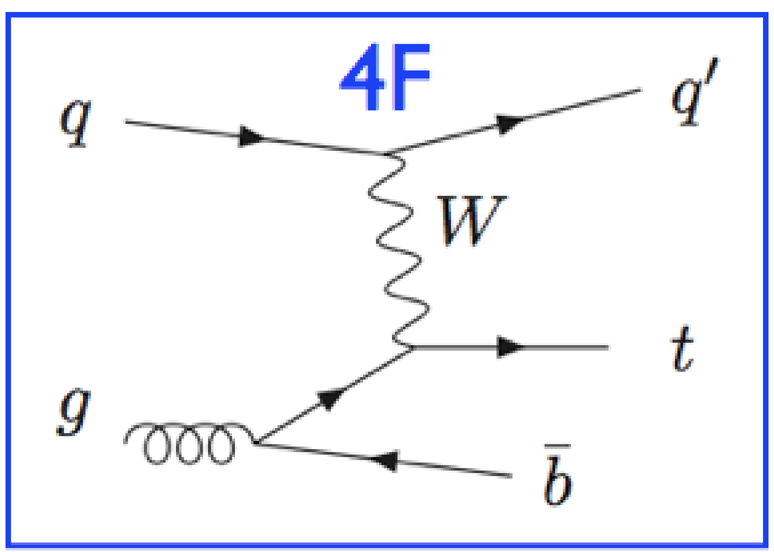


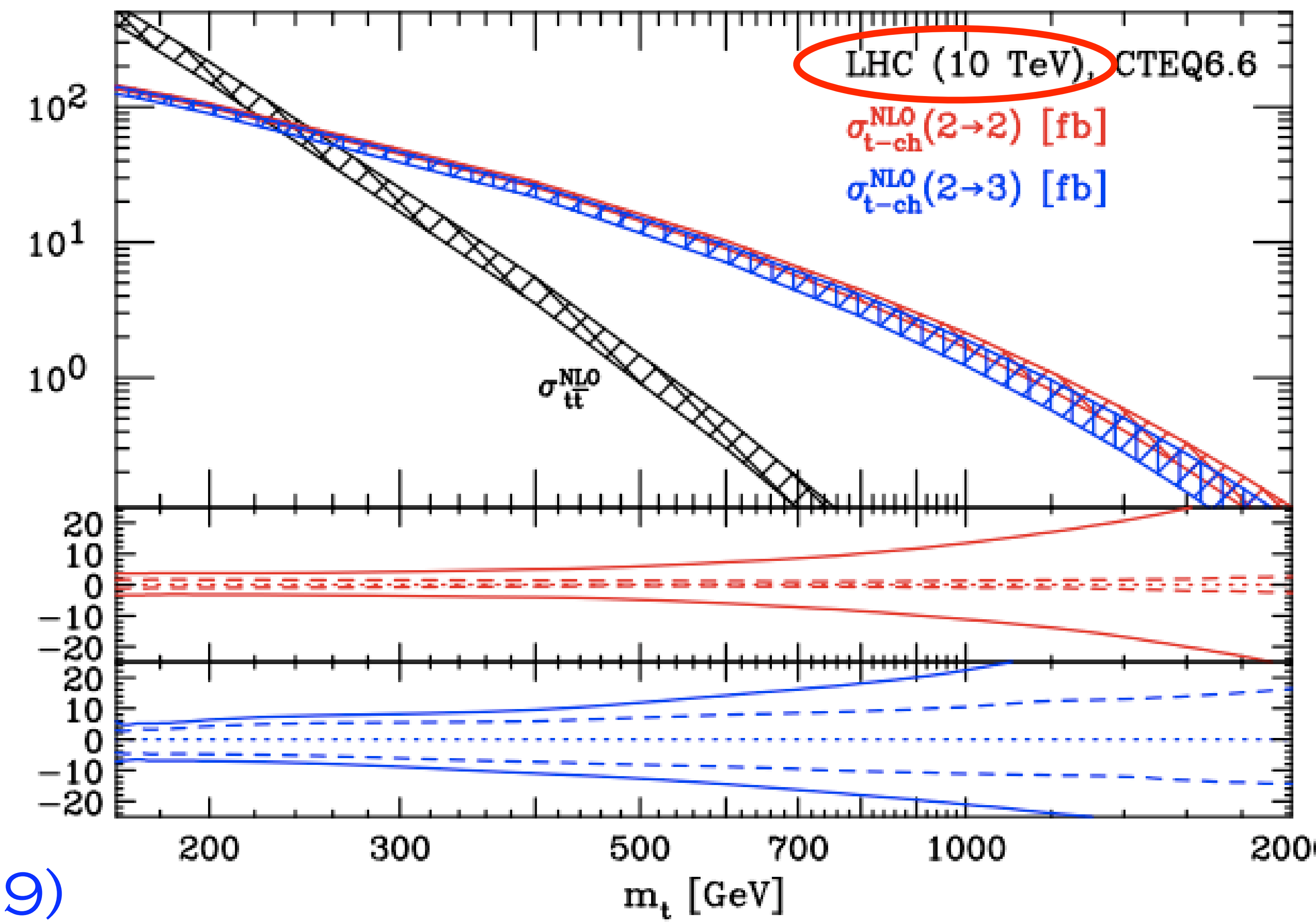
Single top production

- Single top production available in MCFM in both the 4- and 5-flavor schemes.



- Formally equivalent with enough orders.
- At NLO real differences exist. In particular 4F scheme gives access to accompanying b-quark.





JC et al. (2009)

Matching with PS

- No NLO+PS for hadron colliders with 2 or more jets in final state.
  - how do we best use N(NLO) information when no NLO+PS is available?
- Some possible options:

  (a) Use higher orders for overall inclusive normalization only

  ✔ simple to implement, defensible theoretically

  ✘ misses potentially important shape and/or kinematic information

  (b) Split events into jet bins and normalize by best prediction in each bin

  ✔ simple, uses best information, defensible

  ✘ as above + sum of bin cross sections is not a well-defined quantity

  (c) Pick an important distribution and reweight shower to reproduce NLO

  ✔ relatively simple

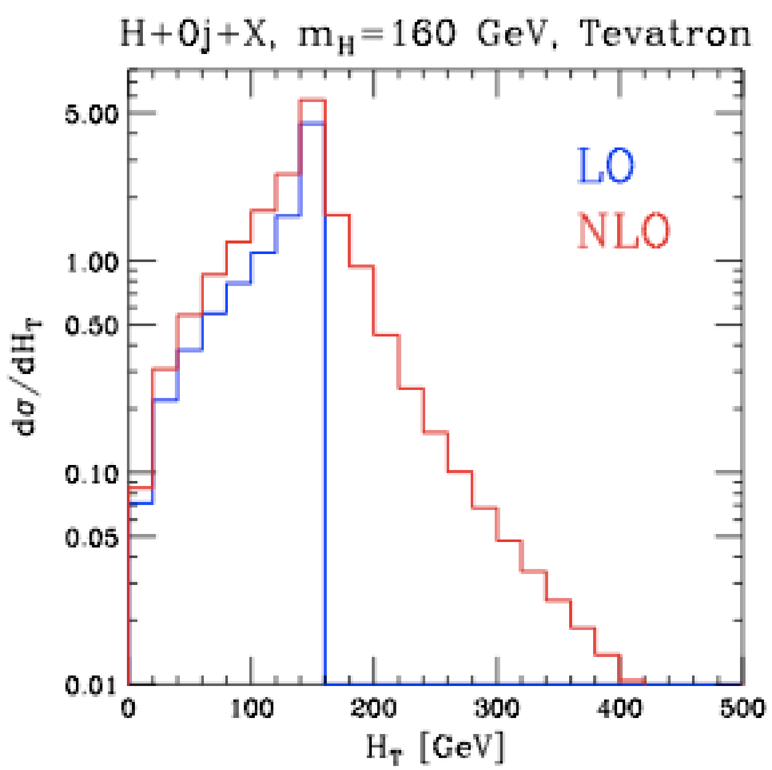
  ✘ throws away some PS shower information; other distributions okay?

  - ....

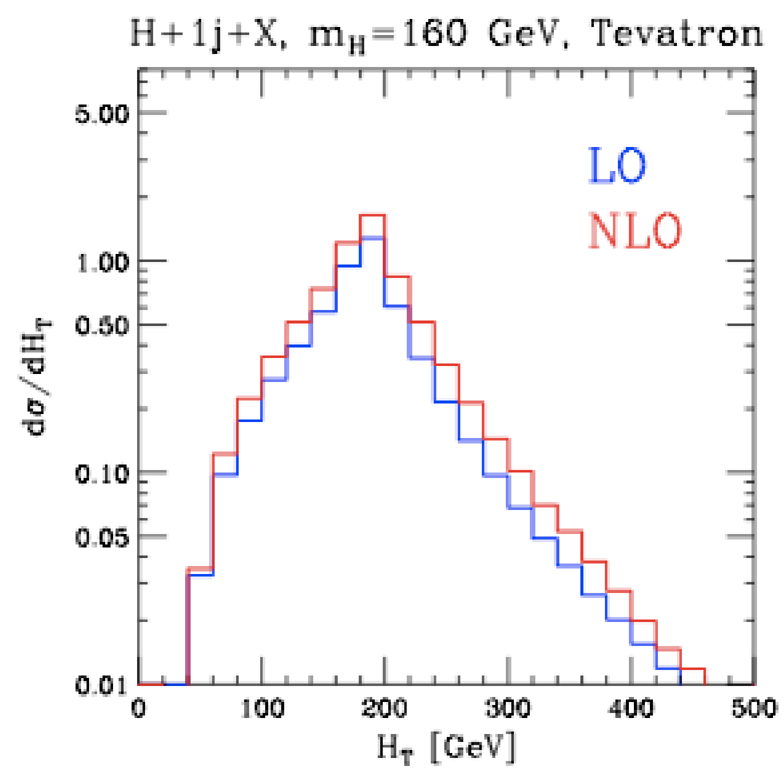
Using NLO: example

- What is really needed is a systematic study, on a case-by-case basis.
- Example: H+jets at the Tevatron.
    - Ongoing work with J. Winter, comparing NLO predictions with SHERPA.

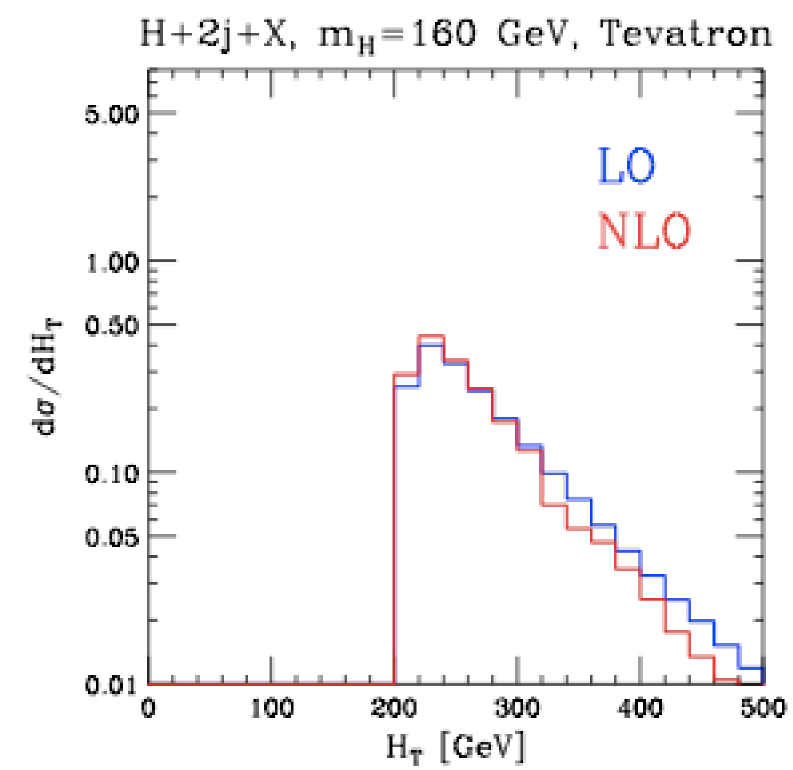
H+X



H+1 jet +X



H+2 jets +X



- Deficiencies of fixed order (esp. LO) clear.
    - How do these compare to SHERPA predictions? Results on the way ...

Summary

- There is a wide array of Monte Carlo tools suitable for LHC analyses:
  - LO: estimates of cross sections, starting point for parton showers.
  - NLO: increasingly available for multijet processes; great advances on the theory side are now turning into LHC phenomenology.
  - NNLO: still very limited; inclusive jets/V+jet/top pairs around the corner.
  - LO + PS + matching: becoming mature, will be even better with tuning on larger data-sets at the Tevatron.
  - NLO + PS: much progress being made, look for V+jet in the near future.
- A variety of signal and background processes are included at NLO in MCFM;
  - new recently: Higgs+2 jets, top pairs with decay, t-channel single top (4F).
  - many processes of interest in early LHC running.
- The MCFM code contains expressions for many virtual matrix elements and could provide a starting point for future implementations in a parton shower.